LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-BOOK-642200

Constraints on 3-D Seismic Models from Global Geodynamic Observables: Implications for the Global Mantle Convective Flow

A. M. Forte, N. A. Simmons, S. P. Grand

August 12, 2013

Treatise on Geophysics, Volume 1 (2nd edition) Seismology & the Structure of the Earth: Theory & Observations

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Constraints on 3-D Seismic Models from Global Geodynamic Observables: Implications for the Global Mantle Convective Flow

Alessandro M. Forte
GEOTOP - Département des Sciences de la Terre et de l'Atmosphère
Université du Québec à Montréal
Montréal, Québec, Canada

Nathan A. Simmons
Atmospheric, Earth and Energy Division
Lawrence Livermore National Laboratory
Livermore, California, USA

Stephen P. Grand
Jackson School of Geosciences
University of Texas at Austin
Austin, Texas, USA

Abstract

Mantle flow models based on seismic tomographic images of 3-D mantle structure may be used to successfully explain a wide range of global surface data related to mantle convection. These tomography-based convection models place strong constraints on viscosity and lateral density variations in Earth's mantle. The convection-related observables which can be modelled include the global free-air gravity anomalies, the dynamic surface and CMB topography, and the present-day tectonic plate motions. The fundamental mantle flow theory needed to interpret these geodynamic observables is reviewed for the case of a self-gravitating, compressible mantle with coupled surface plates. The most important input to the flow models, namely the effective viscosity of the mantle, may be inferred from simultaneous inversions of glacial-isostatic adjustment and mantle convection data. These viscosity inferences are used in numerical calculations of the geodynamic response or kernel functions which provide the theoretical relationship between the surface observables and the mantle density anomalies. Understanding the present-day dynamics of the mantle is critically dependent on knowledge of these internal density perturbations and they are derived here on the basis of a wide selection of recent global tomography models derived from seismic shear velocity (V_S) data. A detailed review of the geodynamic observables predicted on the basis of these V_S tomography models is presented and it is found that all 3-D models derived solely from seismic data provide very similar fits to the convection data. Although the geodynamic fits obtained using the published, purely seismic 3-D mantle models is reasonably good, there remains much room for improvement. The residual misfit, of the order of 50% of the signal in the geodynamic data, raises fundamental questions about the degree to which mantle heterogeneity is

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

resolved by the seismic tomography models or whether the classical fluid mechanical theory employed to model flow in the mantle is adequate. Another outstanding issue in current efforts to understand mantle flow dynamics is the relative importance of the thermal and compositional contributions to mantle heterogeneity. To address these major questions a discussion is presented of recent efforts to successfully reconcile the independent constraints on 3-D mantle structure and flow provided by global seismic and convection-related data sets. The tomographic procedure for jointly inverting these seismic and geodynamic data, including additional constraints from mineral physics data, is found to greatly improve the fits to the convection-related data and this is accomplished in the context of a dominantly thermal origin for the mantle heterogeneity.

1 Introduction

The earliest tomography models (e.g. Dziewonski *et al.* 1977; Dziewonski 1984; Clayton & Comer 1983 – reported in Hager & Clayton 1989; Woodhouse & Dziewonski 1984) exhibited a large-scale three-dimensional (3-D) structure in Earth’s mantle which was reasonably well correlated to the major surface manifestations of mantle convection, namely the large-scale nonhydrostatic geoid (Hager *et al.* 1985) and the large-scale tectonic plate motions (Forte & Peltier, 1987). These early models of mantle heterogeneity thus helped to show that seismic tomography can indeed resolve the lateral variations in mantle structure which drive the convective flow responsible for the ’drift’ of the continents and the large-scale perturbations in Earth’s gravitational field.

Steady progress in global seismic tomographic imaging over the past 15 years has yielded models of 3-D mantle structure which posses significantly improved resolution and reliability (e.g., Ekström & Dziewonski 1998; Ritsema *et al.* 1999; Mégnin & Romanowicz 2000; Masters *et al.* 2000; Boschi & Dziewonski 2000; Grand 2002; Antolik *et al.* 2003; Li *et al.* 2008; Kustowski *et al.* 2008; Simmons *et al.* 2010; Ritsema *et al.* 2011). These improvements are evident in the remarkable accord amongst the images of large-scale mantle structure provided by these different models. Such agreement is encouraging, especially when we consider that these models are obtained on the basis of different data sets, different theoretical treatments of seismic wave propagation and different algorithms for inverting the global seismic data. Helpful reviews of the methods, data sets and current challenges in global seismic imaging may be found in Dziewonski & Romanowicz (2007) and Thurber & Ritsema (2007).

While the analyses and discussion presented below will focus on how seismically inferred 3-D mantle structure may be used to explain geodynamic data and to model mantle flow dynamics, it is important to note that much effort has also been dedicated to the converse approach, namely to infer 3-D mantle structure from geodynamic data. This inverse approach based on geodynamic constraints began at almost the same time as the earliest global tomography inversions. Hager (1984) proposed a mantle flow model based on estimated density anomalies associated with seismically active subducted slabs and used this model to obtain a good match to very long wavelength geoid anomalies. Although subsequent studies, for example by Ricard et al. (1989), attempted a more general inference of mantle heterogeneity using different geodynamic data, the subducted-slab geodynamic models developed by Hager (1984) became prevalent. This led, for example, to a study by Forte & Peltier (1989) which constrained the distribution of subducted slab heterogeneity using other geodynamic observables such as tectonic plate motions and core-mantle boundary topography. The latter study also attempted to use the geodynamic observables to infer the buoyancy forces associated with mantle plumes under mid-ocean ridges.

Tectonic plate motions and long wavelength geoid anomalies have proved to be the most important constraints in developing geodynamic models of 3-D mantle heterogeneity in terms of subducted slabs. The geological history of tectonic plate motions as derived from paleomagnetic data (e.g., Gordon & Jurdy 1986) have been an especially important ingredient in developing models of both past and present-day subducted slab heterogeneity (e.g., Richards & Engebretson 1992; Ricard et al. 1993). These models have been very useful for exploring time-dependent dynamics of the mantle (e.g., Lithgow-Bertelloni & Gurnis 1997; Lithgow-Bertelloni & Richards 1998). Direct comparisons between the long wavelength pattern of subducted slab heterogeneity and the corresponding pattern of heterogeneity derived from seismic tomography have shown good correlations at the longest wavelengths (Richards & Engebretson 1992). When all wavelengths are considered, however, the global correlations between the slab heterogeneity and different tomography models are relatively poor, with correlation coefficients less than 0.3 throughout the mantle (Fig.7 in Lithgow-Bertelloni & Richards 1998).

There are various factors which might explain the less than satisfactory agreement between seismic models of 3-D mantle structure and the reconstructions of mantle heterogeneity in terms of subducted slabs. Firstly, from a purely technical perspective, the evolution of the slab trajectories were not determined in a fluid mechanically consistent manner by numerically solving the advection-diffusion equations. Progress in this direction has been made by Bunge et al. (1998), McNamara & Zhong (2005) and Quéré & Forte (2006) by solving the full set of thermal convection equations with moving tectonic plates as a surface boundary condition. Secondly, the slab models assume that thermally generated heterogeneity dominates in the mantle and this may not be applicable in regions of the mantle with significant compositional heterogeneity (e.g., Forte & Mitrovica 2001; McNamara & Zhong 2005). Perhaps the most important deficiency in these efforts to explain mantle heterogeneity in terms of slab subduction alone is that they do not account for the presence and evolution of hot thermal plumes in the mantle which have been consistently imaged in the global tomography models (e.g., Montelli *et al.* 2004; Nolet *et al.* 2006). A recent appraisal of the origin and importance of mantle plumes in convection models that incorporate 3-D mantle structure derived from seismic tomography may be found in Glisovic *et al.* (2012).

Seismic tomography continues to be the single most important technique for directly inferring the 3-D heterogeneity in the mantle which is associated with the process of thermal convection. The significant advances in global seismic imaging over the past few years provides the underlying motivation for the detailed discussion of the geodynamic implications which will be presented below. A recent detailed review by Becker & Boschi (2002) of the currently available tomography models focussed on a quantitative analysis of similarities and differences between these models. The work presented below will further extend this previous analysis by carrying out calculations of the mantle flow predicted on the basis of these tomography models and examining in detail the extent to which the tomography-based flow models provide a satisfactory explanation for the main convection-related surface observables.

2 Geodynamic Observables and Mantle Flow Theory

In the following, the main global geophysical constraints on mantle structure will be presented. These constraints include the free-air gravity anomalies, the dynamic surface and CMB topography and the tectonic plate motions. The theoretical relationship between these convection-related observables and the internal 3-D structure will be developed in terms of a fluid mechanical model of mantle dynamics.

2.1 Convection-related surface observations

An understanding of thermal convection in the mantle is necessary for explaining a multitude of geophysical and geological processes which we can observe and measure at the surface of the Earth, such as continental drift, earthquakes, mountain building, volcanism, perturbations in Earth's gravitational field, variations in oceanic bathymetry and continental elevation, and long-term changes in global sea-level variations. The principal surface manifestations of mantle convection which have been employed to study the large scale structure and dynamics of the solid Earth are illustrated schematically in figure 1.

The observational constraints on 3-D mantle structure and dynamics which will be considered below are the global free-air gravity anomalies, the dynamic surface and CMB topography and the horizontal divergence of the tectonic plate motions. The Earth's gravitational potential perturbations are usually represented in terms of geoid anomalies however, as pointed out in Forte *et al.* (1994), a more detailed and evenly balanced representation (especially at long wavelengths) of the spectral content is provided by the free-air gravity anomalies shown in figure 2a.

We will employ the term 'dynamic topography' to mean all contributions to Earth's surface topography which arise from density anomalies in the convecting mantle – including the lithosphere. Observational constraints on dynamic topography therefore require an accurate model of crustal heterogeneity in order to remove all isostatic crustal contributions to Earth's measured surface topography. The topographic crustal correction, here based on model CRUST2.0 (Bassin *et al.* 2000), is described in detail by Perry *et al.* (2003) and the resulting dynamic topography is shown in figure 2b.

The tectonic plate velocity field $\mathbf{v}$ may be conveniently summarised in terms of two complimentary scalar fields: the horizontal divergence $\boldsymbol{\nabla}_{\mathbf{H}} \cdot \mathbf{v}$ and the radial vorticity $\hat{\mathbf{r}} \cdot \boldsymbol{\nabla} \times \mathbf{v}$ (Forte & Peltier 1987). The constraint of plate rigidity imposes a linear dependence between these two scalar fields (Forte & Peltier 1991) and hence it suffices that we consider only the plate divergence, shown in figure 2c. (A more detailed discussion of the implications of plate rigidity, in terms of allowed plate motions, is also presented below in section 2.3.7.)

The most robust global constraint on deep mantle density heterogeneity and dynamics (Forte et al. 1995) is currently provided by the dynamical ellipticity of the CMB shown

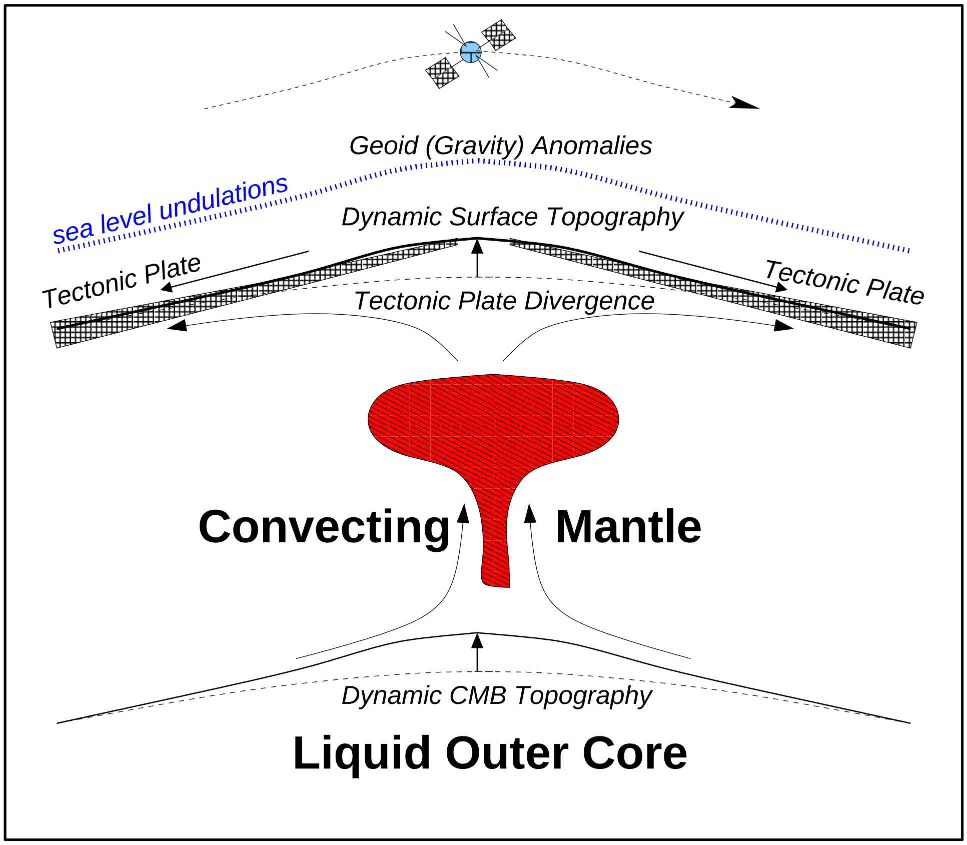
Geoid (Gravity) Anomalies
sea level undulations
Dynamic Surface Topography
Tectonic Plate
Tectonic Plate
Tectonic Plate Divergence
Convecting
Mantle
Dynamic CMB Topography
Liquid Outer Core

Figure 1: Convection-related surface observations. The most direct manifestations of thermal convection in the mantle (represented schematically by the upwelling plume coloured in red) are the surface motions of tectonic plates – which may be summarised by their horizontal divergence field – and the flow induced topography of the solid surface and the core-mantle boundary (CMB). The boundary undulations and the density anomalies in the mantle (due, for example, to the hot upwelling plume) give rise to surface gravitational potential perturbations which may be measured in terms of geoid or gravity anomalies

in figure 2d. The discrepancy between theoretical predictions of the Free Core Nutation (FCN) period (Wahr, 1981) and the value determined from Very Long Baseline Interferometry (Herring et al., 1986), led Gwinn et al. (1986) to conclude that the CMB ellipticity is larger than that implied by theoretical calculations which assumed a rotating Earth in hydrostatic equilibrium. This inference of an 'excess' ellipticity, in which the nonhydrostatic radius of the CMB at the poles is 400 metres less than at the equator, is supported by recent analyses of the FCN period (Mathews et al. 2002).

The most appropriate mathematical basis functions for describing any bounded and continuous function on a spherical surface are the spherical harmonics $Y_\ell^m(\theta, \varphi)$, where position on the spherical surface is defined by colatitude θ and co-longitude φ. We may therefore expand the geodynamic observables in figure 2 in terms of spherical harmonics as follows:

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{+\ell} f_\ell^m \, Y_\ell^m(\theta, \varphi) \tag{1}$$

Where the function $f(\theta, \varphi)$ represents any of the observables (gravity, topography or plate divergence) and the indices ℓ, m which characterise each spherical harmonic are called the *harmonic degree* and *azimuthal order*, respectively. (A useful introduction to spherical harmonic functions may be found in Jackson, 1975.)

The spatial variation of the spherical harmonics is oscillatory in character it may be characterised in terms of an equivalent horizontal wavelength. On a spherical surface of radius r, a spherical harmonic $Y_\ell^m(\theta, \varphi)$ has a characteristic wavelength λ_ℓ given by the following expression:

$$\lambda_\ell = \frac{2\pi r}{\sqrt{\ell(\ell+1)}} \approx \frac{2\pi r}{\ell + \frac{1}{2}} \text{ , valid for } \ell \gg 1 \tag{2}$$

In practise the infinite sum over degree ℓ in (1) is truncated at some finite value ℓ_{max} and for the fields in figure 2 it is $\ell_{max} = 32$. At the Earth's surface (r=6368 km) this is equivalent to a minimum horizontal length scale (or half wavelength) of about 600 km.

Spherical harmonic representations of surface fields can be used to quantify their spectral content in terms of an amplitude spectrum. The amplitude spectrum measures the mean (globally averaged) amplitude of a field at a particular wavelength λ_ℓ (2) and it is defined in terms of a 'degree variance' σ_ℓ as follows:

$$\sigma_\ell = \sqrt{\sum_{m=-\ell}^{+\ell} f_\ell^{m*} f_\ell^m} \tag{3}$$

in which f_ℓ^m are the spherical harmonic coefficients of the surface field (see expression 1) and * denotes complex conjugation.

A spectral description of the spatial correlation between two fields, as a function of harmonic degree or wavelength, is quantified in terms of 'degree correlation' r_ℓ, defined as follows:

$$r_\ell = \frac{\sum_{m=-\ell}^{+\ell} {f_1}_\ell^{m*} {f_2}_\ell^m}{\sqrt{\sum_{m=-\ell}^{+\ell} {f_1}_\ell^{m*} {f_1}_\ell^m} \sqrt{\sum_{m=-\ell}^{+\ell} {f_2}_\ell^{m*} {f_2}_\ell^m}} \tag{4}$$

in which ${f_1}_\ell^m$ and ${f_2}_\ell^m$ are the spherical harmonic coefficients of the two fields.

Degree variances and correlations are very useful spectral characterisations of the amplitude and spatial variation of globally defined surface fields (e.g., O'Connell 1971) and they will be used frequently in the discussions presented below.

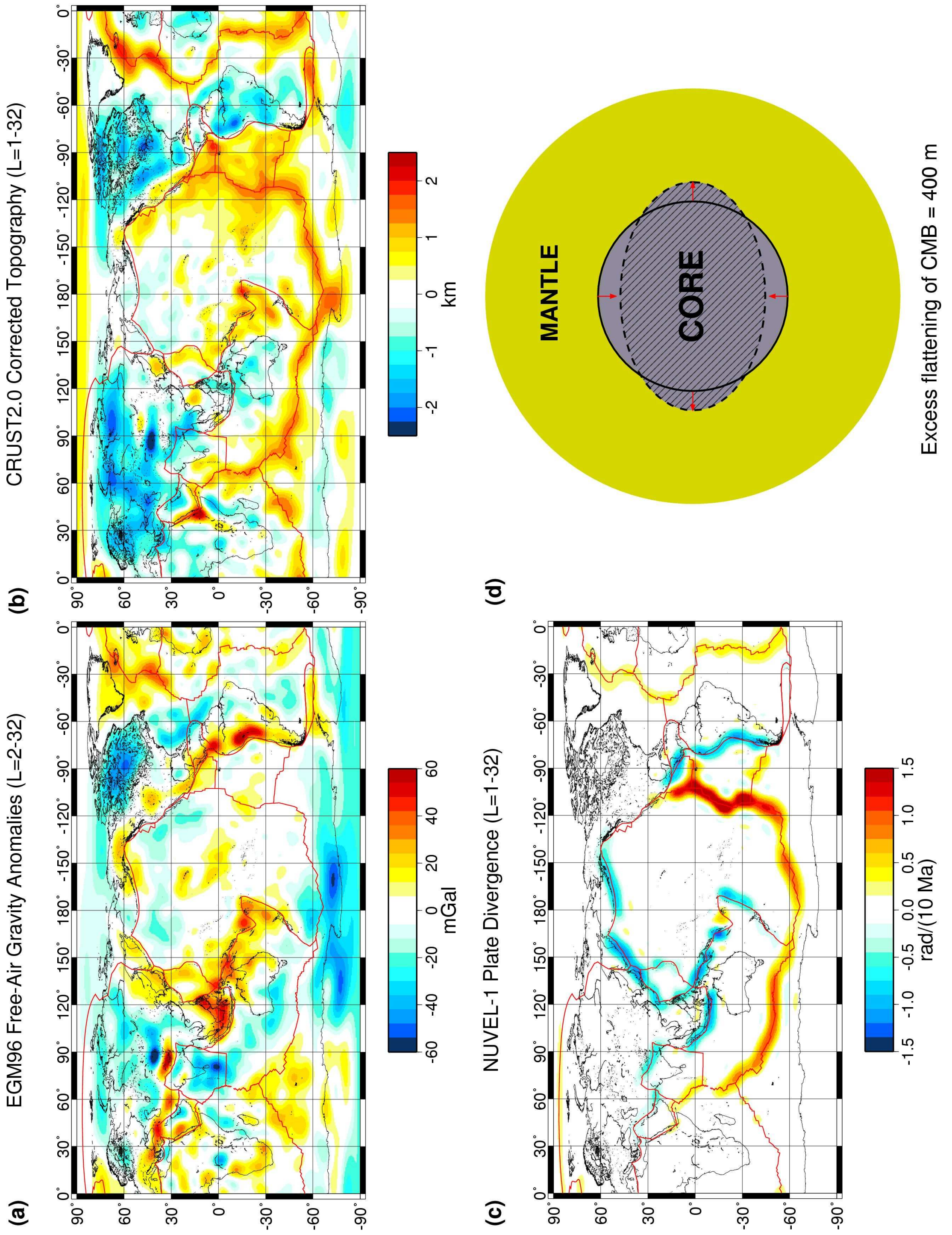
(a)
EGM96 Free-Air Gravity Anomalies (L=2-32)
mGal
-60
-40
-20
0
20
40
60
(b)
CRUST2.0 Corrected Topography (L=1-32)
km
-2
-1
0
1
2
(c)
NUVEL-1 Plate Divergence (L=1-32)
rad/(10 Ma)
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
(d)
MANTLE
CORE
Excess flattening of CMB = 400 m
90°
60°
30°
0°
-30°
-60°
-90°
0°
30°
60°
90°
120°
150°
180°
-150°
-120°
-90°
-60°
-30°
0°

Figure 2: Geodynamic observables. (a) The free-air gravity anomalies derived from the joint geopotential model EGM96 (Lemoine et al. 1998). (b) The dynamic surface topography obtained by removing the topography due to isostatically compensated crustal heterogeneity from Earth's observed topography. The CRUST2.0 (Bassin et al. 2000) crustal heterogeneity model is employed here. (c) The horizontal divergence of the tectonic plate velocities given by the NUVEL-1 model (DeMets et al. 1990). (d) The excess or dynamic CMB ellipticity inferred from core nutation data (Mathews et al. 2002). All fields, with the exception of (d), have been expanded in spherical harmonics up to degree and order 32 (see equation 1).

2.2 Evidence for mantle flow in correlations between internal structure and surface gravity anomalies

The need for a dynamical interpretation of the surface observables (Fig. 2) in terms of very deep-seated convective flow in the mantle is illustrated in Fig. 3 in which long wavelength gravity anomalies are directly compared to seismically inferred lateral heterogeneity in the lower mantle. The large-scale pattern of faster than average shear velocities which lie below the periphery of the Pacific ocean (blue coloured regions in Fig. 3b) has long been interpreted in terms of accumulating lithospheric slabs which have subducted below the margins of the Pacific Ocean since Mesozoic times (Chase 1979; Chase & Sprowl 1983; Richards & Engebretson 1992). These seismic anomalies, which presumably represent positive density anomalies (i.e. cold subducted slabs), are closely correlated with a broad ring-like pattern of negative gravity anomalies around the periphery of the Pacific Ocean (red coloured regions in Fig. 3a).

The correlation between long-wavelength seismic heterogeneity in the lower mantle and surface gravity anomalies was first identified by Dziewonski et al. (1977) and they postulated that a dynamical interpretation in terms of mantle flow could explain the negative sign of the correlation. In a rigid or elastic mantle in which flow is absent, a positive correlation between internal density and surface gravity is expected. However, in a convecting mantle the gravitational signal of internal density anomalies is opposed by the effect of flow-induced topography at the surface and at the CMB and this dynamical balance can lead to the observed negative correlation (e.g. Pekeris 1935; Parsons & Daly 1983; Richards & Hager 1984).

It is instructive to consider a more detailed spectral analysis of the correlation between surface gravity anomalies and seismically inferred lateral heterogeneity at different depths in the mantle. The degree correlations shown in Fig. 4 indicate which depth ranges in the mantle may be important contributors to the different wavelengths of the surface gravity field. The strongest positive correlations appear in the transition-zone region of the mantle (from $\sim$400 to $\sim$800 km depth) and at the longest wavelength (degree 2) they extend down to about 1400 km depth. The positive sign of these correlations indicates that long-wavelength gravity highs overlie similarly large-scale seismically fast regions located at the base of the upper mantle. Negative degree correlations between long wavelength seismic and gravity anomalies are found in the near-surface region, down to about 400 km depth, and in lower mantle below 1400 km depth (as in Fig.3) where the strongest negative correlations are found. These depth-dependent degree correlations, obtained on the basis of long-wavelength tomography models derived more than a decade ago, are robust observations which have been verified using the latest seismic tomographic inferences of 3-D mantle structure (Ricard et al. 2006).

The correlation between mantle heterogeneity and surface gravity anomalies in a convecting mantle has been shown to be a strong function of the depth dependence of the rheological properties of the mantle (e.g. Hager 1984; Ricard et al. 1984; Hager et al. 1985; Forte & Peltier 1987). In particular, as will be discussed further below, the change from positive correlations at the top of the lower mantle to negative correlations at the bottom of lower mantle (Fig. 4) can be interpreted in terms of a significant increase in mean viscosity across the mantle. Other surface observables, such as the tectonic plate motions, also provide important constraints on the depth dependence of mantle rheology (e.g. Ricard & Vigny 1989; Forte et al. 1991). The geodynamic surface observables also provide direct constraints on the 3-D distribution of density anomalies in the mantle and hence these observables provide a fundamentally important and independent means of evaluating the extent to which seismic tomography models successfully resolve the lateral heterogeneity in the mantle which is associated with mantle convection.

In the next section the dynamical link between lateral heterogeneity in the mantle and convection-related surface observables will be developed in the theoretical framework of a fluid mechanical model of viscous flow in the mantle. The reader wishing to skip over this detailed mathematical treatment of the mantle flow theory is encouraged to jump directly to sections 2.4 and 2.5, where the viscous response of the mantle and current mantle viscosity inferences, respectively, are summarised.

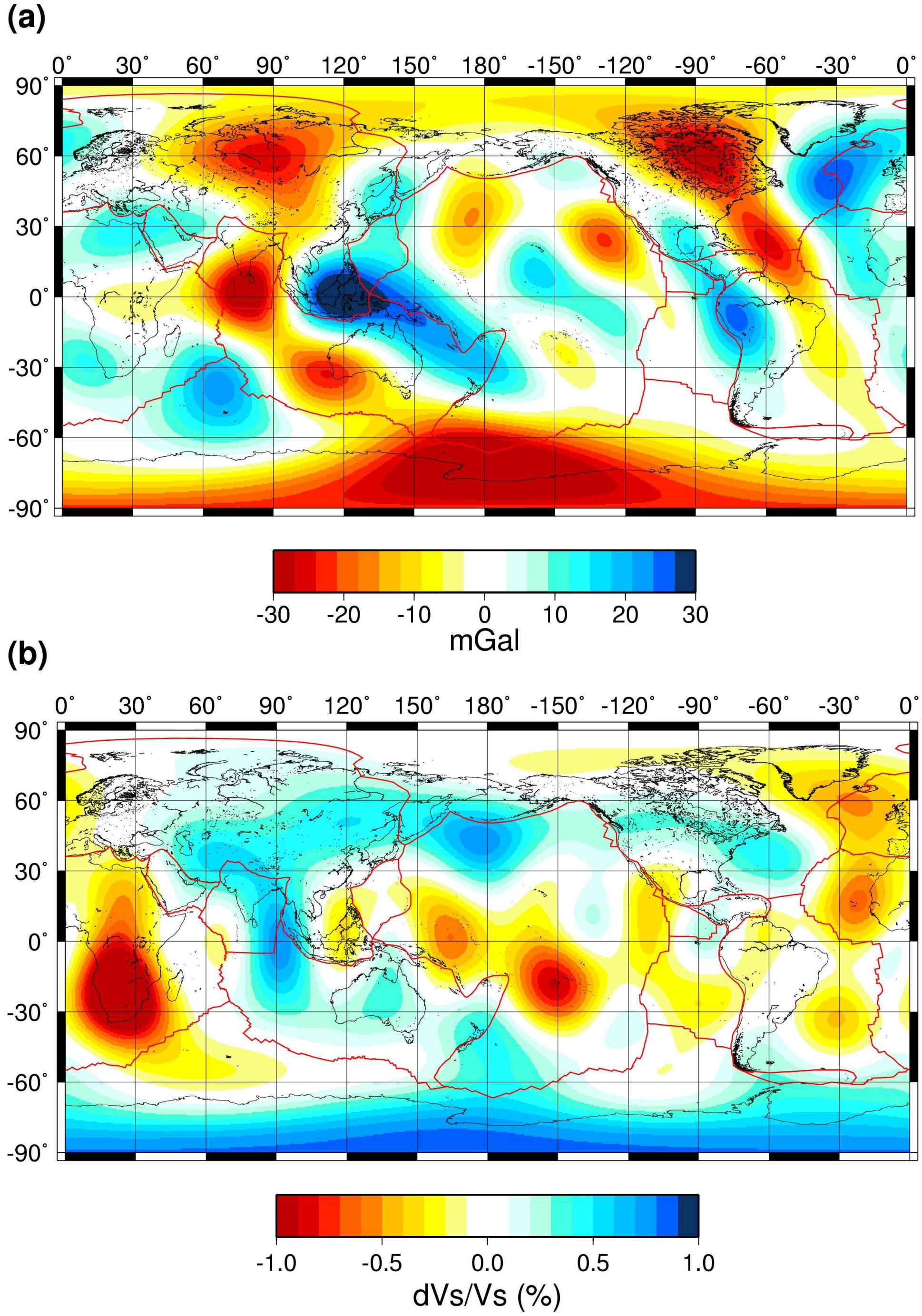
(a)
0˚
30˚
60˚
90˚
120˚
150˚
180˚
-150˚
-120˚
-90˚
-60˚
-30˚
0˚
90˚
60˚
30˚
0˚
-30˚
-60˚
-90˚
-30
-20
-10
0
10
20
30
mGal
(b)
0˚
30˚
60˚
90˚
120˚
150˚
180˚
-150˚
-120˚
-90˚
-60˚
-30˚
0˚
90˚
60˚
30˚
0˚
-30˚
-60˚
-90˚
-1.0
-0.5
0.0
0.5
1.0
dVs/Vs (%)

Figure 3: Surface gravity anomalies and deep-mantle heterogeneity. (a) Long wavelength free-air gravity anomalies derived from the joint geopotential model EGM96 (Lemoine et al. 1998). (b) Long wavelength seismic shear velocity anomalies at a depth of 2100 km derived from the tomography model of Grand (2002). Both the gravity and seismic anomaly fields are synthesized from spherical harmonics in the degree range $\ell = 1 - 8$.

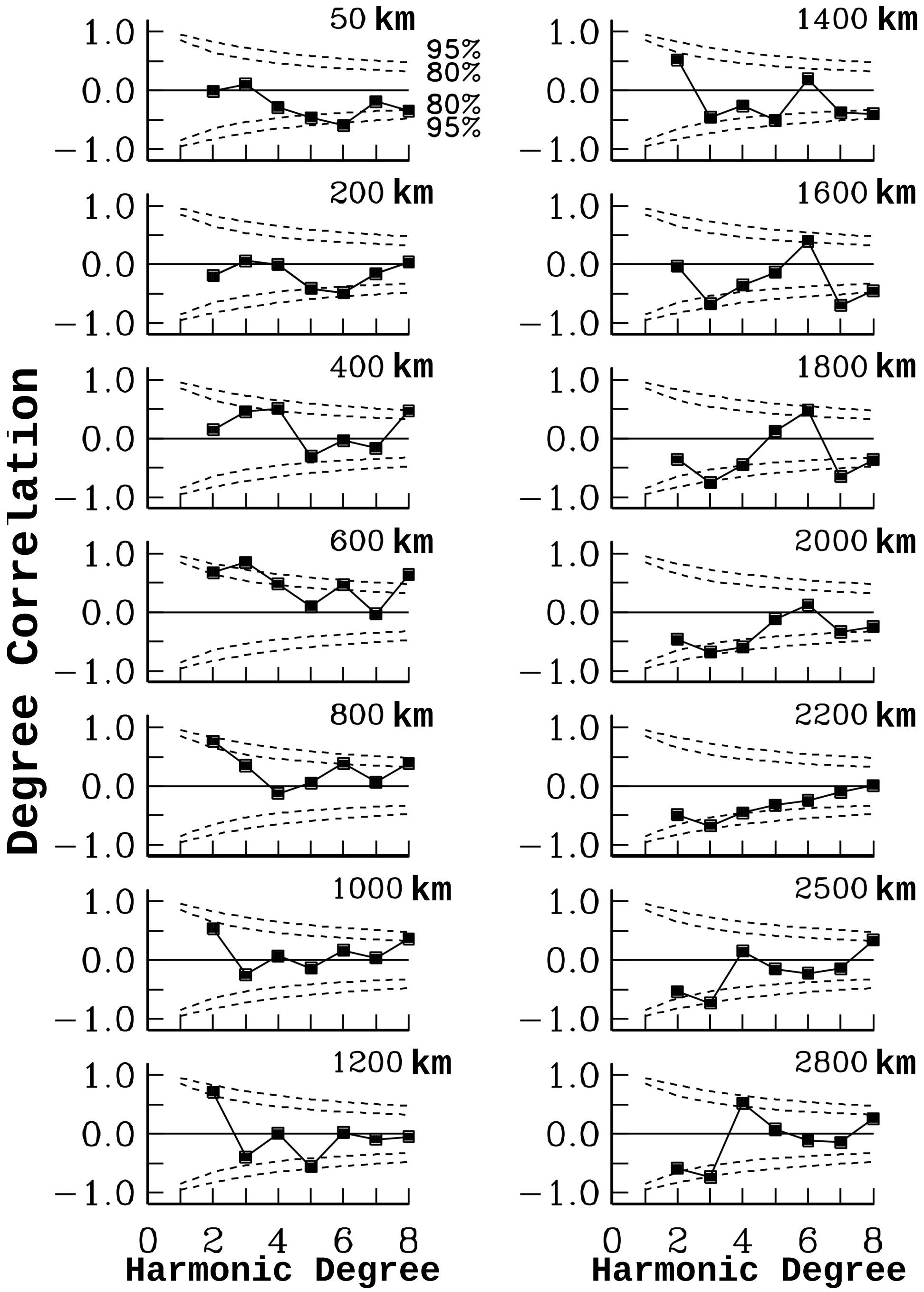

Figure 4: Depth-dependent correlations between 3-D mantle structure and gravity anomalies. Each frame shows the degree-correlations between surface gravity anomalies and lateral variations in seismic shear velocity at a particular depth in the mantle. The seismic anomalies are derived from an earlier long-wavelength tomography model *SH8/U4L8* (Forte et al. 1993a). Figure is adapted from Forte et al. (1993a).

2.3 Fluid mechanical modelling of viscous mantle flow

2.3.1 Governing equations

Laboratory and geological evidence of the ability of mantle rocks to creep indefinitely over geological time scales is understood in terms of the existence of atomic-scale defects in the lattice of crystal grains (e.g. Nicolas & Poirier 1976). If the ambient mantle temperature is sufficiently high, the imposition of nonhydrostatic stresses causes the lattice defects to propagate and thus allows the mantle rocks to creep or 'flow' slowly. This process may be characterised in terms of a single parameter, namely an effective viscosity (e.g. Weertman & Weertman 1975; Weertman 1978).

The characterisation of the long-term creep properties of the mantle in terms of an effective viscosity allows us to model the slow flow of the mantle with the conventional hydrodynamic field equations. The hydrodynamic field equations (Landau & Lifshitz 1959) which express the principles of conservation of mass and momentum are

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{u}) = 0 \tag{5}$$

$$\rho \frac{d\mathbf{u}}{dt} = \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} + \rho \mathbf{g} \tag{6}$$

in which $\mathbf{u}$ is the velocity field, $d/dt = \partial/\partial t + \mathbf{u} \cdot \boldsymbol{\nabla}$ is the total (Lagrangian) time derivative, $\boldsymbol{\sigma}$ is the stress tensor, g is the gravitational acceleration and ρ is the density.

We may represent the gravity field $\mathbf{g}$ in terms of a gravitational potential:

$$\mathbf{g} = \boldsymbol{\nabla} \phi \tag{7}$$

Notice that the sign convention adopted here is opposite to that generally adopted in classical physics, where a *negative* gradient of the potential is used. With this sign convention, Poisson's equation for the relationship between the gravity potential and density is:

$$\Delta \phi = -4\pi \mathrm{G} \rho \qquad (\Delta \equiv \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}) \tag{8}$$

An explicit expression for the stress tensor is given by

$$\sigma_{ij} = -P\delta_{ij} + \tau_{ij} \tag{9}$$

$$\text{with } \tau_{ij} = \eta \left(u_{i,j} + u_{j,i} - \frac{2}{3}\delta_{ij} u_{k,k} \right) + \lambda \delta_{ij} u_{k,k} \tag{10}$$

where τ_{ij} is the viscous stress tensor, δ_{ij} is the identity tensor, P is the total pressure, η and λ are the *isotropic* viscosity coefficients, and $u_{i,j} = \partial u_i / \partial x_j$ represents the derivative of the velocity components i with respect to the coordinate direction j.

A number of simplifications are possible when applying the above field equations to the problem of flow in the mantle:

- Since mantle rocks creep much slower than the acoustic velocity in the mantle, we can safely ignore the term $\partial \rho / \partial t$ in the conservation of mass equation (5) and we thus have the following anelastic-liquid approximation:

$$\boldsymbol{\nabla} \cdot (\rho \mathbf{u}) = 0 \tag{11}$$

This simplification will explicitly rule out acoustic waves as solutions of the flow equations.

- The second viscosity coefficient λ in the viscous stress tensor τ_{ij} describes the dissipation associated with change in fluid volume (density). This volume dissipation may be neglected if the changes in fluid volume occur on time scales which are much longer than for molecular relaxation processes (Landau & Lifshitz 1959), and this is certainly true for mantle flow. Therefore, the viscous stress tensor τ_{ij} will be purely deviatoric:

$$\tau_{ij} = \eta \left(u_{i,j} + u_{j,i} - \frac{2}{3}\delta_{ij} u_{k,k} \right) \tag{12}$$

- We will further assume a Newtonian (i.e. linear) rheology, in which the mantle viscosity η is not a function of stress or strain-rate. This assumption is not necessarily appropriate for the mantle (particularly in high stress regions, such as subduction zones) but it will greatly simplify the mathematical resolution of the flow equations which is presented below.

- A major simplification, from a purely mathematical standpoint, derives from the assumption that mantle viscosity varies only with radius. Although this assumption will significantly simplify the mathematical treatment of the flow theory it may have important physical implications for the dynamics of buoyancy induced flow in the mantle (e.g., Richards & Hager 1989). These implications have been examined previously in the context of global-scale flow in 3-D spherical geometry (e.g., Ricard et al. 1988; Martinec *et al.* 1993; Zhang & Christensen 1993; Forte & Peltier 1994; Čadek & Fleitout 2003; Moucha *et al.* 2007) and they will be discussed below.

- A fundamental physical simplification arises from the very large value of the effective viscosity in the mantle. We may non-dimensionalise the conservation of momentum equation with the following variable transformations:

$$\left. \begin{array}{c} (x, y, z) = (d\, x, d\, y, d\, z) \\[2ex] t = (d^2/\kappa_o)\, t \\[2ex] \rho = \rho_o\, \rho\ ;\ \mathbf{g} = \mathbf{g}_o\, \mathbf{g}\ ;\ \eta = \eta_o\, \eta \end{array} \right\} \tag{13}$$

where the original variables are on the left and the non-dimensional ones are on the right. Quantities d, κ_o, ρ_o, η_o, g_o are characteristic scales for length, thermal diffusivity, density, viscosity, and gravitational acceleration, respectively, in the mantle. The use, in (13), of a thermal diffusion time scale

$$t_\kappa = d^2/\kappa_o \tag{14}$$

is appropriate since we are dealing with mantle flow arising from thermal convection. The stress tensor and pressure have the same physical units (see expression 9), and hence their non-dimensional transformation is:

$$P = (\rho_o g_o d)\ P \quad \text{and} \quad \boldsymbol{\tau} = (\rho_o g_o d)\ \boldsymbol{\tau} \tag{15}$$

where, again, the original quantities are on the left and the non-dimensional ones on the right. By virtue of the constitutive relation (12), and expression (15), we obtain the following non-dimensional transformation for the flow variable:

$$\mathbf{u} = \frac{\rho_o g_o d^2}{\eta_o}\, \mathbf{u} \tag{16}$$

or equivalently,

$$\mathbf{u} = g_o\, t_\nu\, \mathbf{u} \tag{17}$$

where $t_\nu = d^2/\nu_o$ is the time scale for viscous diffusion of momentum and $\nu_o = \eta_o/\rho_o$ is the kinematic viscosity. Finally, substitution of expressions (9, 13-17) into (6) yields the following non-dimensional momentum conservation equation:

$$\rho \left[\frac{1}{Pr} \frac{\partial \mathbf{u}}{\partial t} + \left(\frac{g_o}{d} \right) t_\nu^2\, \mathbf{u} \cdot \boldsymbol{\nabla} \mathbf{u} \right] = \boldsymbol{\nabla} \cdot \boldsymbol{\tau} - \boldsymbol{\nabla} P + \rho \mathbf{g} \tag{18}$$

in which $Pr = t_\kappa / t_\nu$ is the Prandtl number, which characterises the ratio of temperature and momentum diffusion time scales.

To assess the importance of the terms on the left hand side of equation (18), we assume the following values for the scaling variables: $d = 3 \times 10^6$ m (for whole-mantle flow), $\rho_o = 3.3 \times 10^3$ kg/m^3, $g_o = 10$ m/s^2, $\kappa_o = 1.5 \times 10^{-6}$ m^2/s, $\eta_o = 10^{21}$ Pa s, and we thus obtain:

$$t_\nu = 3 \times 10^{-5} \text{ s}; \quad t_\kappa = 6 \times 10^{18} \text{ s}; \quad \frac{g_o}{d} = 3.3 \times 10^{-6} \text{ s}^{-2}; \quad Pr = 2 \times 10^{23} \tag{19}$$

The key quantity here is the characteristic time for momentum diffusion t_ν which is very small (i.e. mantle flow will come to a complete halt in much less than a millisecond if all buoyancy forces are suddenly removed).

The vanishingly small momentum diffusion time, and hence the very large Prandtl number, implies that the inertial forces ($\rho d\mathbf{u}/dt$) in the momentum conservation equation are completely negligible. Therefore, using expression (7), the momentum conservation equation (6) simplifies to:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\tau} - \boldsymbol{\nabla} P + \rho \boldsymbol{\nabla} \phi = \mathbf{0} \tag{20}$$

Equation (20) shows that in the absence of inertia, there must at all times be a balance between the buoyancy forces ρg and the forces of viscous disspation described by $\boldsymbol{\nabla} \cdot \boldsymbol{\tau}$. In other words, any changes in internal buoyancy forces will instantly produce changes in fluid flow: this is a consequence of the essentially instantaneous diffusion of momentum in the mantle.

We may define an idealised hydrostatic reference state for the mantle, which corresponds to the absence of any internal flow or deformation (i.e. $\mathbf{u} = \mathbf{0}$). In this situation the deviatoric stress field $\boldsymbol{\tau}$ vanishes and the momentum conservation equation (6) reduces to:

$$-\boldsymbol{\nabla} P_o + \rho_o \boldsymbol{\nabla} \phi_o = \mathbf{0} \tag{21}$$

in which P_o, ρ_o, and ϕ_o are the pressure, density and gravity potentials in the hydrostatic state. Poisson's equation (8) for a hydrostatic planet is:

$$\Delta \phi_o = -4\pi \mathrm{G} \rho_o \tag{22}$$

We assume that in a dynamic mantle, with a non-vanishing mantle flow $\mathbf{u}$, the pressure, density and gravity potentials will be perturbed as follows:

$$P = P_o + P_1 \qquad \rho = \rho_o + \rho_1 \qquad \phi = \phi_o + \phi_1 \tag{23}$$

in which all perturbations are assumed to be small, that is:

$$\left|\frac{P_1}{P_o}\right| \ll 1 \qquad \left|\frac{\rho_1}{\rho_o}\right| \ll 1 \qquad \left|\frac{\phi_1}{\phi_o}\right| \ll 1$$

If we now substitute the perturbed variables (23) into the equations of mass and momentum conservation (11–20) and Poisson's equation (8), and then substract out the hydrostatic reference equations (21–22), we finally obtain the following set of first-order accurate, perturbed equations for mantle flow dynamics:

$$\textit{mass conservation} \qquad \boldsymbol{\nabla} \cdot (\rho_o \mathbf{u}) = 0 \tag{24}$$

$$\textit{momentum conservation} \qquad \boldsymbol{\nabla} \cdot \boldsymbol{\tau} - \boldsymbol{\nabla} P_1 + \rho_o \boldsymbol{\nabla} \phi_1 + \rho_1 \boldsymbol{\nabla} \phi_o = \mathbf{0} \tag{25}$$

$$\textit{gravity} \qquad \Delta \phi_1 = -4\pi \mathrm{G} \rho_1 \tag{26}$$

Notice in equation (25), that in addition to the driving buoyancy forces ($\rho_1 \boldsymbol{\nabla} \phi_o$), there also exist self-gravitational loads ($\rho_o \boldsymbol{\nabla} \phi_1$) due to the perturbed gravity field. The above equations must be supplemented by the linear relationship (12) between stress and strain-rate, which is valid for an isotropic rheology:

$$\textit{Newtonian constitutive equation} \qquad \boldsymbol{\tau} = \eta \left(\overrightarrow{\boldsymbol{\nabla}} \mathbf{u} + \mathbf{u} \overleftarrow{\boldsymbol{\nabla}} - \tfrac{2}{3} \, \mathbf{I} \, \boldsymbol{\nabla} \cdot \mathbf{u} \right) \tag{27}$$

2.3.2 Spectral treatment of the mantle flow equations

A classical technique for solving the dynamical flow equations in 3-D spherical geometry is the spectral Green function method, in which all flow variables are expressed in terms of spherical harmonic basis functions introduced in equation (1) above (e.g., Hager & O'Connell 1981; Richards & Hager 1984; Ricard *et al.* 1984; Forte & Peltier 1987,1991). The spectral Green functions provide a very convenient mathematical description of the instantaneous relationship between the mantle density anomalies and the viscous flow field, as well as all surface manifestations of the internal flow dynamics.

The spectral method is employed here for solving the coupled equations of mass, momentum and gravity conservation (eqs. 24-26), supplemented by the viscous constitutive relation (27). The method follows that initially developed by Forte & Peltier (1991) for gravitationally consistent compressible flow in spherical geometry. Other treatments of compressible mantle flow in spherical geometry have been presented by Corrieu *et al.* (1995), Panasyuk *et al.* (1996), and Defraigne (1997). An overview of the method is presented below and we refer to Forte & Peltier (1991), Forte & Woodward (1997a) and Forte (2000) for complete details.

One begins by rewriting equations (24-27), in the following Cartesian tensor form:

$$\left.\begin{array}{c} u_{k,k} = -\frac{\dot{\rho_o}}{\rho_o} u_r \\ \sigma_{ij,j} + \rho_o \left(\phi_1\right)_{,i} - \rho_1 \mathrm{g}_o \hat{\boldsymbol{r}} = 0 \\ \sigma_{ij} = -P_1 \delta_{ij} + \eta \left(u_{i,j} + u_{j,i} - \frac{2}{3} \delta_{ij} u_{k,k}\right) \\ \left(\phi_1\right)_{,kk} = -4\pi \mathrm{G} \rho_1 \end{array}\right\} \tag{28}$$

in which $\dot{\rho_o} = d\rho_o/dr$, $u_r = \hat{\boldsymbol{r}} \cdot \mathbf{u}$, and $(\phi_o)_{,i} = -\mathrm{g}_o \hat{\boldsymbol{r}}$. It should also be noted that in these equations the total stress tensor $\boldsymbol{\sigma}$ is used, rather than the deviatoric stress $\boldsymbol{\tau}$ used in (25).

The determination of a solution to the system of tensor equations (28) in spherical geometry may be greatly simplified by using an elegant mathematical technique described by Phinney & Burridge (1973). Following their technique, one introduces a new coordinate system defined by the following complex basis vectors in spherical geometry:

$$\left.\begin{array}{rcl} \hat{\boldsymbol{e}}_- & = & \frac{1}{\sqrt{2}} \left(\hat{\boldsymbol{\vartheta}} - \imath\, \hat{\boldsymbol{\varphi}}\right) \\ \hat{\boldsymbol{e}}_0 & = & \hat{\boldsymbol{r}} \\ \hat{\boldsymbol{e}}_+ & = & -\frac{1}{\sqrt{2}} \left(\hat{\boldsymbol{\vartheta}} + \imath\, \hat{\boldsymbol{\varphi}}\right) \end{array}\right\} \tag{29}$$

in which $\imath = \sqrt{-1}$, and $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\vartheta}}$, $\hat{\boldsymbol{\varphi}}$ are the unit basis vectors for the standard spherical polar coordinate system.

Following Phinney & Burridge (1973), all the tensors appearing in the original system (28) are rotated into the coordinate system defined by (29), thereby yielding the following covariant tensor form of the dynamical equations:

$$\left.\begin{array}{c} u^{\alpha,\beta}\, e_{\alpha\beta} = -\frac{\dot{\rho_o}}{\rho_o} u^0 \\ \sigma^{\alpha\beta,\gamma}\, e_{\beta\gamma} + \rho_o \left(\phi_1\right)^{,\alpha} - \rho_1 \mathrm{g}_o\, \delta^\alpha_0 = 0 \\ \sigma^{\alpha\beta} = -P_1\, e^{\alpha\beta} + \eta \left(u^{\alpha,\beta} + u^{\beta,\alpha}\right) - \frac{2}{3}\eta \left(u^{\delta,\gamma}\, e_{\delta\gamma}\right) e^{\alpha\beta} \\ \left(\phi_1\right)^{,\alpha\beta}\, e_{\alpha\beta} = -4\pi \mathrm{G} \rho_1 \end{array}\right\} \tag{30}$$

in which the Greek indices denote the coordinate directions in system (29) and therefore range over the values $(-1, 0, +1)$. The quantities $e^{\alpha\beta}$, $e_{\alpha\beta}$, and δ^α_β are the contravariant, covariant, and mixed tensor representations of the Cartesian identity tensor δ_{ij}.

The velocity (u^α) and stress ($\sigma^{\alpha\beta}$) components are expanded in terms of the generalised spherical functions $Y_\ell^{N\,m}$ (Phinney & Burridge 1973) such that:

$$\begin{aligned}
u^-(r,\theta,\phi) &= \sum_{\ell,m} U_\ell^{-\,m}(r)\, Y_\ell^{-1\,m}(\theta,\phi), \\
u^0(r,\theta,\phi) &= \sum_{\ell,m} U_\ell^{0\,m}(r)\, Y_\ell^{0\,m}(\theta,\phi), \\
u^+(r,\theta,\phi) &= \sum_{\ell,m} U_\ell^{+\,m}(r)\, Y_\ell^{+1\,m}(\theta,\phi), \qquad (31)\\
\sigma^{\alpha\beta}(r,\theta,\phi) &= \sum_{\ell,m} T_\ell^{\alpha\beta m}(r)\, Y_\ell^{(\alpha+\beta)\,m}(\theta,\phi) \quad (\alpha,\beta=-1,0,+1). \qquad (32)
\end{aligned}$$

All scalar fields involved in the governing equations (30) are expanded in terms of ordinary spherical harmonics. For example,

$$\begin{aligned}
\rho_1(r,\theta,\phi) &= \sum_{\ell,m} (\rho_1)_\ell^m(r)\, Y_\ell^m(\theta,\phi)\ , \\
\phi_1(r,\theta,\phi) &= \sum_{\ell,m} (\phi_1)_\ell^m(r)\, Y_\ell^m(\theta,\phi)\ , \qquad (33)
\end{aligned}$$

where $Y_\ell^m \equiv Y_\ell^{0\,m}$.

We can simplify subsequent numerical computations by non-dimensionalise all relevant physical variables using the following transformations:

$$\left.\begin{array}{llll}
r = d\,r, & \mathrm{g}_o(r) = \mathrm{g}_o\,\mathrm{g}(r), & \eta(r) = \eta_o\,\eta(r), & \rho_1 = (\Delta\rho)\ \rho_1, \\
T^{\alpha\beta} = (\Delta\rho\,\mathrm{g}_o\,d)\ T^{\alpha\beta}, & & U^\alpha = \left(\frac{\Delta\rho\,\mathrm{g}_o\,d^2}{\eta_o}\right) U^\alpha, & \phi_1 = \left(\frac{4\pi \mathrm{G} R_o}{2\ell+1}\,\Delta\rho\, d\right)\ \phi_1
\end{array}\right\} \qquad (34)$$

in which the original variables are on the left of each equation and the non-dimensional variables are on the right. The scaling quantities we have used are defined as follows:

$d = 2888\ km$ ≡ radial thickness of mantle
$\mathrm{g}_o = 9.82\ m/s^2$ ≡ mean surface gravitational acceleration
$\Delta\rho = 0.1\ Mg/m^3$ ≡ characteristic subducted slab density anomaly
$\eta_o = 10^{21}\ Pa\,s$ ≡ *Haskell* [1935] reference value
$R_o = 6371\ km$ ≡ mean surface radius of Earth

Substituting expressions (31-32, 33) into the flow equations (30), and using the nondimensionalization scheme (34) as well as the orthogonality properties of the generalised spherical harmonics (as in Phinney & Burridge 1973), yields two independent systems of flow equations. The first system, governing a poloidal geometry of flow, consists of the following six coupled, first-order, ordinary differential equations:

$$\frac{d}{dr}\begin{pmatrix} U^0 \\ U^P \\ T^0 \\ T^P \\ \phi_1 \\ g_1 \end{pmatrix} = \begin{pmatrix}
-\frac{[2+r(\dot{\rho}_o/\rho_o)]}{r} & \frac{\Omega_1}{r} & 0 & 0 & 0 & 0 \\
-\frac{2\Omega_1}{r} & \frac{1}{r} & 0 & \frac{\eta_o}{\eta} & 0 & 0 \\
\frac{4[3+r(\dot{\rho}_o/\rho_o)]}{r^2}\frac{\eta}{\eta_o} & -\frac{6\Omega_1}{r^2}\frac{\eta}{\eta_o} & 0 & \frac{\Omega_1}{r} & \frac{3}{2\ell+1}\frac{\dot{\rho}_o}{\overline{\rho}} & 0 \\
-\frac{4\Omega_1[3+r(\dot{\rho}_o/\rho_o)]}{r^2}\frac{\eta}{\eta_o} & \frac{2\left[\Omega_2^2+3\Omega_1^2\right]}{r^2}\frac{\eta}{\eta_o} & -\frac{2\Omega_1}{r} & -\frac{3}{r} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & \frac{\ell(\ell+1)}{r^2} & -\frac{2}{r}
\end{pmatrix}\begin{pmatrix} U^0 \\ U^P \\ T^0 \\ T^P \\ \phi_1 \\ g_1 \end{pmatrix} +$$

$$\begin{pmatrix} 0 \\ 0 \\ g(r) \\ 0 \\ 0 \\ (-2\ell + 1)\ \frac{d}{R_o} \end{pmatrix} \frac{\rho_1(r)}{\Delta\rho} \tag{35}$$

in which $\dot{\rho_o} = d\rho_o/dr$, $\Omega_1 = [\ell(\ell+1)/2]^{1/2}$, $\Omega_2 = [(\ell-1)(\ell+2)/2]^{1/2}$, and $\overline{\rho}$ is Earth's mean density (5.5143×10^3 kg/m^3).

The second system, governing a toroidal geometry of flow, consists of the following two coupled, first-order, ordinary differential equations:

$$\frac{d}{dr}\begin{pmatrix} U^T \\ T^T \end{pmatrix} = \begin{pmatrix} \frac{1}{r} & \frac{1}{\eta} \\ \frac{2\,\Omega_2^2\,\eta}{r^2} & -\frac{3}{r} \end{pmatrix} \begin{pmatrix} U^T \\ T^T \end{pmatrix} \tag{36}$$

The flow, stress and gravity variables in the poloidal and toroidal flow systems (35,36) are all dependent on the same harmonic degree ℓ and order m (this dependence has been dropped for notational convenience). The dependence of these variables on the harmonic coefficients of flow(31), stress (32), and gravitational potential (33) are:

$$\begin{aligned} U^P(r) &= U^+(r) + U^-(r) \\ U^T(r) &= U^+(r) - U^-(r) \\ T^0 &= T^{00}(r) + \frac{3}{2\ell+1}\frac{\rho_0(r)}{\overline{\rho}}\,\phi_1(r) \\ T^P &= T^{0+}(r) + T^{0-}(r) \\ T^T &= T^{0+}(r) - T^{0-}(r)\ . \end{aligned} \tag{37}$$

2.3.3 Internal boundary conditions

The determination of a unique solution of the system of flow equations (35,36) derived in the preceding section requires the specification of appropriate boundary conditions at the top and bottom surfaces of the mantle and internal matching conditions at all material interfaces within the mantle. We will consider the latter first since the surface boundary conditions can be obtained as a special case of the internal matching conditions.

The need for internal boundary conditions in solving for mantle flow is evident upon inspection of the depth variation of mantle density $\rho_o(r)$ (see Fig. 5) given by the PREM seismic reference model (Dziewonski & Anderson 1981). The PREM density profile is characterised by two major jumps at depths of 400 km and 670 km. These density jumps have long been interpreted as manifestations of phase-change boundaries (e.g., Jeanloz & Thompson 1983) which presumably also affect the depth variation of viscosity at these depths (e.g., Sammis et al. 1977).

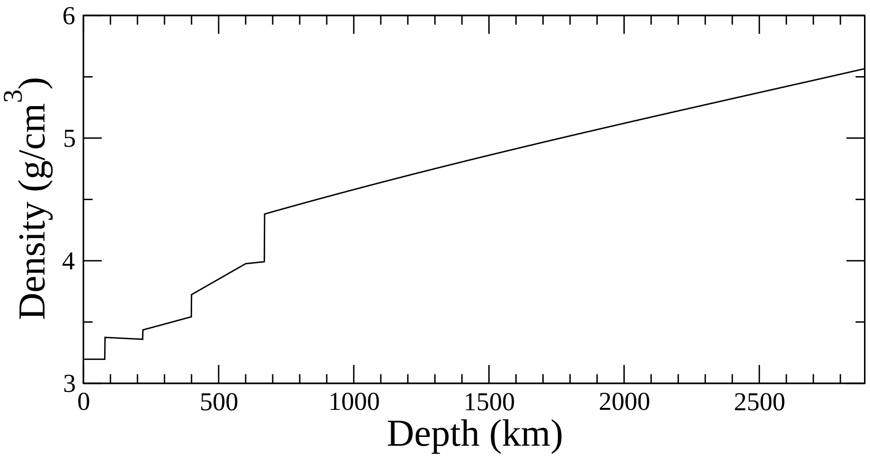

Figure 5: Depth dependent mean density in the mantle from the PREM seismic reference model (Dziewonski & Anderson 1981).

We now derive the matching conditions which must be satisfied when characteristic properties of the mantle, in particular density, viscosity and chemical composition, change very rapidly across phase-change or chemical horizons in the mantle. We will approximate such rapid changes as mathematical discontinuities. Denoting the mean (i.e. horizontally averaged) location of the internal material boundary as $r = r_i$, geographic variations in the radial location (i.e., deflections) of the boundary are then represented by:

$$r = r_i + \delta r_i \,, \quad \text{where we assume} \quad \left| \frac{\delta r_i}{r_i} \right| \ll 1 \tag{38}$$

A discontinuous change in density across an internal boundary implies that:

$$\left(\frac{d\rho_o}{dr} \right)_{r=r_i} = \left[\rho_o(r_i^+) - \rho_o(r_i^-) \right] \delta(r - r_i) \tag{39}$$

By virtue of this last expression, the density perturbation due to the boundary deflection δr_i is given by the following sheet-mass anomaly:

$$\delta\rho_i = - \left(\frac{d\rho_o}{dr} \right)_{r=r_i} \delta r_i = -[\rho_o(r_i^+) - \rho_o(r_i^-)] \, \delta r_i \, \delta(r - r_i) \tag{40}$$

This last expression shows that an internal boundary deflection gives rise to a mantle buoyancy source which may be approximated by a delta-function load and the corresponding density perturbation is, in non-dimensional terms:

$$\frac{\rho_1(r)}{\Delta\rho} = - \frac{[\rho_o(r_i^+) - \rho_o(r_i^-)]}{\Delta\rho} \, \frac{\delta r_i}{d} \, \delta(r - r_i) \tag{41}$$

This expression shows that from a purely fluid mechanical perspective, the dynamical effect of a deformed phase-change boundary is indistinguishable from integrated buoyancy forces elsewhere in the mantle. If the material interface instead corresponds to a chemical discontinuity in the mantle, the fluid-mechanical buoyancy effect is also the same. The main distinction between the deflection-induced buoyancy of phase-change and chemical discontinuities is that the former is controlled by thermodynamics (i.e., the Clapeyron slope) whereas the the latter is due to the change in intrinsic chemical buoyancy across the interface.

The matching condition for the radial component of the flow field is obtained by integrating the first row in system (35) from $r_i^- = r_i - \epsilon$ to $r_i^+ = r_i + \epsilon$, taking the limit $\epsilon \to 0$. In terms of the radial flow variable U^0, this yields the following mass conservation equation:

$$\rho_o(r_i^+)U^0(r_i^+) = \rho_o(r_i^-)U^0(r_i^-) \tag{42}$$

where $\rho_o(r_i^+)$ and $\rho_o(r_i^-)$ are the mantle density immediately above and below the material boundary, respectively. If the internal boundary corresponds to a chemical discontinuity, then we must impose a zero radial velocity condition :

$$U^0(r_i^+) = 0 = U^0(r_i^-) \tag{43}$$

The matching condition for the tangential mantle flow velocity is obtained by integrating the second equation in (35), which yields the following continuity of the tangential flow variable U^P:

$$U^P(r_i^+) = U^P(r_i^+) \tag{44}$$

This continuity of tangential flow also applies to the case of a internal chemical boundary.

The internal matching condition for the radial stress T^0 across a deformed material interface is obtained by integrating the third row in system (35), in which we substitute result (41) into the buoyancy force term $\rho_1/\Delta\rho$ in expression (35) thereby yielding:

$$\begin{aligned} T^0(r_i^+) - T^0(r_i^-) = \Delta_u(r_i) + \frac{3}{2\ell+1}\left[\frac{\rho_o(r_i^+) - \rho_o(r_i^-)}{\overline{\rho}}\right]\phi_1(r_i) \\ - \frac{[\rho_o(r_i^+) - \rho_o(r_i^-)]}{\Delta\rho}\frac{\delta r_i}{d}\,\mathrm{g}(r_i) \end{aligned} \tag{45}$$

in which the term

$$\Delta_u(r_i) = -\frac{2}{r_i}\left(\frac{\eta(r_i^+) + \eta(r_i^-)}{\eta_o}\right)\left[U^0(r_i^+) - U^0(r_i^-)\right] \tag{46}$$

arises from the discontinuous change in radial flow (see expression 42) due to a density jump across an internal (phase-change) boundary. This term vanishes in the case of a chemical discontinuity since expression (43) then applies. The matching condition for the tangential stress T^P is obtained by integrating the fourth equation in (35):

$$T^P(r_i^+) - T^P(r_i^-) = -\Omega_1\Delta_u(r_i) \tag{47}$$

The term Δ_u gives rise to an apparent discontinuity in tangential stress and this is again a consequence of our mathematical treatment of a discontinuous change in radial flow across a material interface. The significance of this apparent jump in tangential stress was first noted by Corrieu et al. (1995) and a subsequent analysis by Forte & Woodward (1997a) determined that effect of this discontinuity was negligible. In the case of a chemical discontinuity the term Δ_u vanishes and the tangential stress is then continuous across the interface.

Integration of the fifth row in (35) across the material interface located at mean radius r_i yields the following matching condition for the perturbed gravitational potential:

$$\phi_1(r_i^+) = \phi_1(r_i^-) \tag{48}$$

The matching condition for the perturbed gravitational acceleration is obtained by substituting expression (41) into the buoyancy force term $\rho_1/\Delta\rho$ in (35) and then integrating the sixth row of this system, yielding the following non-dimensional expression:

$$g_1(r_i^+) - g_1(r_i^-) = (2\ell + 1)\left(\frac{d}{R_o}\right)\frac{[\rho_o(r_i^+) - \rho_o(r_i^-)]}{\Delta\rho}\,\frac{\delta r_i}{d} \tag{49}$$

in which $g_1 = d\phi_1/dr$.

The internal matching conditions for the toroidal components of mantle flow are readily obtained following the same procedure employed for the poloidal flow. Integration of the first row in system (36) yields the following continuity of toroidal flow across any (phase-change or chemical) interface:

$$U^T(r_i^+) = U^T(r_i^-) \tag{50}$$

Similarly, an integration of the second row in (36) yields the following continuity of tangential stress across any internal material interface:

$$T^T(r_i^+) = T^T(r_i^-) \tag{51}$$

.

2.3.4 Boundary conditions at Earth's solid surface

It will be assumed that the spherically symmetric, hydrostatic reference Earth model is overlain by a global ocean layer which is 3 km thick, as in PREM (Dziewonski & Anderson 1981). PREM's crust and seismic lithosphere (LID) are combined into a single mechanical layer containing the same total mass as the two PREM layers assuming, for simplicity, that both layers deform and move together in response the buoyancy driven flow in the mantle. The top surface of the combined crust-lithosphere is located at radius $r = 6368$ km and the base is located at radius 6291 km (i.e., at a depth of 80 km below the surface of the global ocean layer). This redefined lithosphere has mass density of $3.2\ Mg/m^3$.

In view of the very complicated mechanical and rheological properties of the crust and underlying lithosphere, in particular the tectonic plates, it is clear that the flow theory developed here with a purely depth-dependent viscosity cannot provide an adequate representation of the near-surface dynamics. An approximate treatment of the effect of surface tectonic plates is presented below in section (2.3.7). In the meantime, we will consider here two different boundary conditions of relevance at the solid-surface: *free-slip* and *no-slip*.

Earth's bounding surface at $r = a = 6368$ km is a chemical (compositional) boundary across which there can be no flow and hence the matching condition (43) is applicable:

$$U^0(a^-) = 0, \;\; \text{for both } \textit{free-slip} \;\&\; \textit{no-slip} \tag{52}$$

where a^- denotes the radial location $r = a - \epsilon$. We can, to first-order accuracy, ignore surface deflections (i.e. topography) in specifying the boundary conditions on the flow.

The condition for the surface tangential flow $U^P(a^-)$ in the lithosphere is:

$$U^P(a^-) = \begin{cases} U^P(a^-) & \text{to be determined, for } \textit{free-slip} \\ 0 & \text{for } \textit{no-slip} \end{cases} \tag{53}$$

In the case of the surface toroidal flow, the boundary condition is:

$$U^T(a^-) = V^T \tag{54}$$

in which V^T is determined from the coupling of surface poloidal and toroidal flows due to the presence of rigid surface tectonic plates. The mathematical formulation of this coupling and an explicit expression for V^T is presented below in section (2.3.7).

We can apply the radial stress T^0 matching condition in (45) to the deformed surface boundary which, by virtue of conditions (52), becomes:

$$T^0(a^-) - T^0(a^+) = \frac{3}{2\ell+1}\left[\frac{\Delta\rho_a}{\overline{\rho}}\right]\phi_1(a^-) - \left[\frac{\Delta\rho_a}{\Delta\rho}\right]\frac{\delta a}{d}\,\mathrm{g}(a) \tag{55}$$

where we have defined the density jump across the solid surface:

$$\Delta\rho_a = \rho_o(a^-) - \rho_o(a^+) = 3.2 - 1.0\; Mg/m^3 = 2.2\; Mg/m^3 \tag{56}$$

and where we have also invoked the universally valid condition (48) for the vertical continuity of the perturbed gravitational potential.

In expression (55), $T^0(a^+)$ corresponds to the radial stress in the global ocean layer. If we assume that the viscosity in the global ocean layer is negligible (i.e., $\eta/\eta_o \to 0$), then the second row in system (35) yields:

$$T^P(r) = 0\,, \;\text{throughout the ocean layer} \tag{57}$$

and therefore, by virtue of this result, the fourth row in system (35) yields:

$$T^0(r) = 0\,, \;\text{throughout the ocean layer} \tag{58}$$

Substituting result (58) into expression (55) yields the desired radial-stress boundary condition at the surface:

$$T^0(a^-) = \frac{3}{2\ell+1}\left[\frac{\Delta\rho_a}{\overline{\rho}}\right]\phi_1(a^-) - \left[\frac{\Delta\rho_a}{\Delta\rho}\right]\frac{\delta a}{d}\,\mathrm{g}(a) \;\; \text{valid for } \textit{free-slip \& no-slip} \tag{59}$$

in which δa is the flow-induced vertical deflection of the solid surface (i.e., dynamic surface topography).

The condition for the surface poloidal tangential stress $T^P(a^-)$ is as follows:

$$T^P(a^-) = \begin{cases} 0 & \text{for } \textit{free-slip} \\ T^P(a^-) & \text{to be determined, for } \textit{no-slip} \end{cases} \tag{60}$$

and the condition for the surface toroidal tangential stress is:

$$T^T(a^-) = T^T(a^-) \quad \text{to be determined} \tag{61}$$

On the basis of the general result (48), and using result (49), the surface matching conditions for the perturbed gravitational potential and acceleration are:

$$\phi_1(a^-) \;=\; \phi_1(a^+) \tag{62}$$

$$g_1(a^-) \;=\; g_1(a^+) + (2\ell+1)\left(\frac{d}{R_o}\right)\left[\frac{\Delta\rho_a}{\Delta\rho}\right]\frac{\delta a}{d} \tag{63}$$

The ocean-layer potential and gravity fields $\phi_1(a^+)$ and $g_1(a^+)$, respectively, are not independent of each other and, as shown in Forte & Peltier (1991), they are both related to the perturbed potential at the surface of the global ocean layer, $\phi_1(r = R_o)$:

$$\phi_1(a^+) = P_\ell\, \phi_1(R_o) \tag{64}$$

$$g_1(a^+) = G_\ell\, \phi_1(R_o) \tag{65}$$

in which the ocean-layer response functions P_ℓ and G_ℓ are as follows:

$$P_\ell \;=\; \left(\frac{R_o}{a}\right)^{\ell+1} - \frac{3}{2\ell+1}\frac{\rho_w}{\bar\rho}\left[\left(\frac{R_o}{a}\right)^{\ell+2} - \left(\frac{a}{R_o}\right)^{\ell-1}\right] \tag{66}$$

$$\begin{aligned} G_\ell \;=\; & -(\ell+1)\frac{d}{R_o}\left(\frac{R_o}{a}\right)^{\ell+2} \\ & + \frac{3}{2\ell+1}\frac{\rho_w}{\bar\rho}\frac{d}{R_o}\left[(\ell+1)\left(\frac{R_o}{a}\right)^{\ell+3} + \ell\left(\frac{a}{R_o}\right)^{\ell-2}\right] \end{aligned} \tag{67}$$

in which $\rho_w \equiv \rho(a^+) = 1\ Mg/m^3$ is the density of the global ocean layer.

Substitution of results (64-65) into expressions (62-63) yields the complete surface boundary conditions for the gravitational variables:

$$\phi_1(a^-) = P_\ell\, \phi_1(R_o)\,, \ \text{valid for } \textit{free-slip \& no-slip} \tag{68}$$

$$g_1(a^-) = G_\ell\, \phi_1(R_o) + (2\ell+1)\left(\frac{d}{R_o}\right)\left[\frac{\Delta\rho_a}{\Delta\rho}\right]\frac{\delta a}{d}\,, \ \text{valid for } \textit{free-slip \& no-slip} \tag{69}$$

It should be noted that expression (68) should also be substituted into the radial stress condition (59).

The complete set of free-slip and no-slip surface boundary conditions, in terms of the poloidal-flow vector $\mathbf{v}(r) = [U^0(r),\ U^P(r),\ T^0(r),\ T^P(r),\ \phi_1(r),\ g_1(r)]^{Tr}$ (where Tr denotes transposition) employed in system (35), are:

Free-slip:

$$\mathbf{v}(a^-) = U^P(a^-)\,\mathbf{y}_1 + \phi_1(R_o)\,\mathbf{y}_2 + \frac{\Delta\rho_a}{\Delta\rho}\frac{\delta a}{d}\,\mathbf{y}_3 \tag{70}$$

in which the surface basis vectors are:

$$\mathbf{y}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{y}_2 = \begin{pmatrix} 0 \\ 0 \\ \frac{3}{2\ell+1}\frac{\Delta\rho_a}{\bar{\rho}}\,P_\ell \\ 0 \\ P_\ell \\ G_\ell \end{pmatrix} \quad \mathbf{y}_3 = \begin{pmatrix} 0 \\ 0 \\ -\mathrm{g}(a) \\ 0 \\ 0 \\ (2\ell+1)\left(\frac{d}{R_o}\right) \end{pmatrix} \tag{71}$$

No-slip:

$$\mathbf{v}(a^-) = T^P(a^-)\,\mathbf{y}'_1 + \phi_1(R_o)\,\mathbf{y}_2 + \frac{\Delta\rho_a}{\Delta\rho}\frac{\delta a}{d}\,\mathbf{y}_3 \tag{72}$$

in which the surface basis vectors are:

$$\mathbf{y}'_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and } \mathbf{y}_2\ ,\ \mathbf{y}_3 \text{ are defined in (71)} \tag{73}$$

The surface boundary conditions (54,61) for the toroidal flow vector $\mathbf{u}(r) = [U^T(r),\ T^T(r)]^{Tr}$ which is governed by system (36) are:

$$\mathbf{u}(a^-) = V^T\,\mathbf{z}_1 + T^T(a^-)\,\mathbf{z}_2 \tag{74}$$

in which the two surface basis vectors are:

$$\mathbf{z}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \mathbf{z}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{75}$$

2.3.5 Boundary conditions at CMB

The derivation of the boundary conditions which apply at the core-mantle boundary (CMB), located at mean radius $r = b = 3480$ km, is almost identical to the derivation for the surface boundary conditions in the preceding section. The liquid outer core is regarded as having negligible viscosity relative to the mantle and hence the CMB is treated as a purely free-slip boundary. The only difference concerns the application of gravitational matching conditions at $r = b$, since we must now deal with the interaction between a deformed CMB and a compressible, hydrostatic core.

A detailed treatment of the gravitational perturbations maintained in a hydrostatic core is presented in Forte & Peltier (1991), where it is shown that perturbed gravitational acceleration at the top of the core (i.e., immediately below the CMB) is determined by the perturbed potential at the bottom of the mantle (i.e., immediately above the CMB), as follows:

$$g_1(b^-) = R_\ell \, \phi_1(b^+)$$

where b^- denotes the radial location $r = b - \epsilon$ (i.e., bottom side of the CMB) and b^+ denotes $r = b + \epsilon$ (i.e., top side of the CMB). R_ℓ is a numerically determined coefficient which is obtained on the basis of the compressible density profile throughout the core (Forte & Peltier 1991).

The complete set of free-slip CMB boundary conditions in terms of the poloidal-flow vector $\mathbf{v}(r) = [U^0(r),\ U^p(r),\ T^0(r),\ T^P(r),\ \phi_1(r),\ g_1(r)]^{Tr}$ in system (35) is:

$$\mathbf{v}(b^+) = U^P(b^+)\,\mathbf{x}_1 + \phi_1(b^+)\,\mathbf{x}_2 + \frac{\Delta\rho_b}{\Delta\rho}\frac{\delta b}{d}\,\mathbf{x}_3 \tag{76}$$

in which the CMB basis vectors are:

$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 0 \\ \frac{3}{2\ell+1}\frac{\Delta\rho_b}{\overline{\rho}} \\ 0 \\ 1 \\ R_\ell \end{pmatrix} \quad \mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ -\mathrm{g}(b) \\ 0 \\ 0 \\ (2\ell+1)\left(\frac{d}{R_o}\right) \end{pmatrix} \tag{77}$$

where $\Delta\rho_b = \rho_o(b^+) - \rho(b^-) = -4.434\ Mg/m^3$ is the density jump across the CMB and δb is the deflection (i.e., dynamic topography) of the CMB.

The free-slip CMB boundary condition for the toroidal-flow vector $\mathbf{u}(r) = [U^T(r),\ T^T(r)]^{Tr}$ in system (36) is:

$$\mathbf{u}(b^+) = U^T(b^+)\,\mathbf{w}_1 \tag{78}$$

in which the CMB basis vector is:

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{79}$$

2.3.6 Determining viscous flow Green functions

The solution to the system of flow equations (35) may expressed in terms of Green functions which relate the poloidal flow velocity and the stress tensor at an arbitrary radius $r = r_0$ to a delta-function density load $\rho_1(r) = \delta(r - r')$ at any other radius $r = r'$.

For arbitrarily complex density (e.g., as in PREM) and viscosity profiles, the poloidal-flow system of equations (35) must be integrated numerically. When $r \neq r'$, this linear system of equations is homogeneous:

$$\frac{d}{dr}\mathbf{v}(r) = \mathbf{M}_P(r)\,\mathbf{v}(r)\,, (\textit{when}\ \ r \neq r') \tag{80}$$

in which $\mathbf{M}_P(r)$ is the 6×6 matrix appearing in system (35). We can propagate each of the surface boundary vectors $\mathbf{y}_i$ ($i = 1, 2, 3$) in (71) or (73), by numerically integrating (80) from the surface ($r = a^-$) downward to $r = r'$, stopping along the way at all internal material boundaries ($r = r_i$) where we apply the internal matching conditions described previously (section 2.3.3). We can similarly propagate each of the CMB basis vectors $\mathbf{x}_i$ ($i = 1, 2, 3$) in (77) from the CMB ($r = b^+$) upward to $r = r'$. The basic procedure is summarised schematically in Fig. 6.

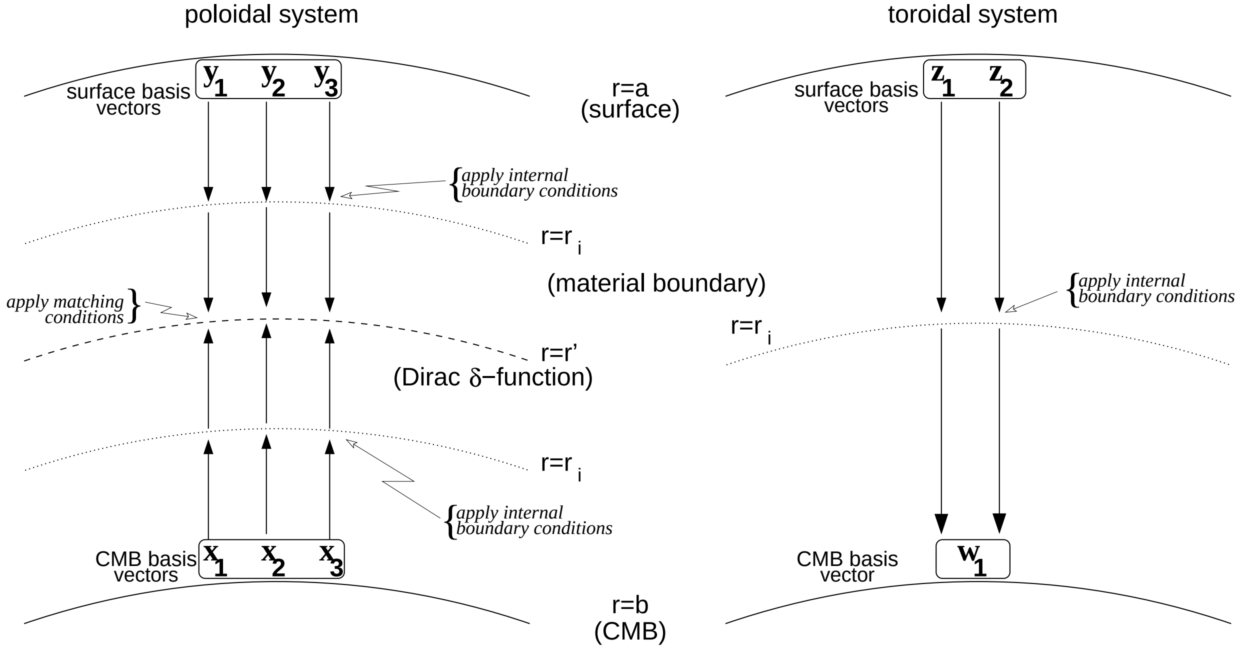
poloidal system
toroidal system
surface basis vectors
y1 y2 y3
r=a
(surface)
surface basis vectors
z1 z2
apply internal boundary conditions
r=r i
(material boundary)
apply matching conditions
apply internal boundary conditions
r=r i
r=r'
(Dirac δ−function)
r=r i
apply internal boundary conditions
CMB basis vectors
x1 x2 x3
CMB basis vector
w1
r=b
(CMB)

Figure 6: Numerical integration of poloidal-flow system of equations (35) (*left*) and toroidal-flow system of equations (36) (*right*). The poloidal-flow basis vectors $\mathbf{y}_i$ and $\mathbf{x}_i$ are defined in expressions (71,73,77). The toroidal-flow basis vectors are defined in expressions (75,79).

At the location $r = r'$ of the delta-function load we apply the matching conditions (42, 44, 45, 47, 48, 49). For each of these matching conditions we employ the corresponding components of the poloidal-flow vectors $\mathbf{v}(r'_+)$ and $\mathbf{v}(r'_-)$, which are obtained from expressions (70), (72) and (76):

$$\mathbf{v}(r'_+) = \left\{ \begin{matrix} U^P(a^-)\,\mathbf{y}_1(r'_+) \\ T^P(a^-)\,\mathbf{y}'_1(r'_+) \end{matrix} \right\} + \phi_1(R_o)\,\mathbf{y}_2(r'_+) + \frac{\Delta\rho_a}{\Delta\rho}\frac{\delta a}{d}\,\mathbf{y}_3(r'_+) \tag{81}$$

and

$$\mathbf{v}(r'_-) = U^P(b^+)\,\mathbf{x}_1(r'_-) + \phi_1(b^+)\,\mathbf{x}_2(r'_-) + \frac{\Delta\rho_b}{\Delta\rho}\frac{\delta b}{d}\,\mathbf{x}_3(r'_-) \tag{82}$$

in which the surface and CMB basis vectors, $\mathbf{y}_i(r'_+)$ and $\mathbf{x}_i(r'_-)$, have been obtained by the numerical integration of (80), as outlined above. The application of the six matching conditions then yields the following system of equations:

$$\mathbf{A}\,\mathbf{p} = \boldsymbol{\delta_p} \tag{83}$$

in which

$$\mathbf{p} = \begin{pmatrix} U^P(a^-)\ or\ T^P(a^-) \\ \phi_1(R_o) \\ \frac{\Delta\rho_a}{\Delta\rho}\frac{\delta a}{d} \\ U^P(b^+) \\ \phi_1(b^+) \\ \frac{\Delta\rho_b}{\Delta\rho}\frac{\delta b}{d} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\delta_p} = \begin{pmatrix} 0 \\ 0 \\ \mathrm{g}(r') \\ 0 \\ 0 \\ -(2\ell+1)\frac{d}{R_o} \end{pmatrix} \tag{84}$$

Each column of the 6×6 matrix $\mathbf{A}$ in (83) involves the corresponding basis vectors $\mathbf{y}_i(r'_+)$ and $\mathbf{x}_i(r'_-)$ in (81-82).

For any given position $r = r'$ of the delta-function load we obtain a system of equations given by (83) which can be simply solved to find the unknown vector $\mathbf{p}$. The elements of $\mathbf{p}$ define the poloidal-flow impulse response (kernel) functions of the mantle which are discussed below in section (2.4).

The toroidal-flow system (36) may be written as the following homogeneous equation:

$$\frac{d}{dr}\mathbf{u}(r) = \mathbf{M}_T(r)\,\mathbf{u}(r) \tag{85}$$

in which $\mathbf{M}_T(r)$ is the 2×2 matrix appearing in system (36). We can propagate each of the surface boundary vectors $\mathbf{z}_i$ $(i = 1, 2)$ in (75), by numerically integrating (85) from the surface $(r = a^-)$ downward to $r = b^+$, stopping along the way at all internal material boundaries $(r = r_i)$ where we apply the internal matching conditions (50,51). The basic procedure is summarised schematically in Fig. 6. The resulting flow vector $\mathbf{u}(b^+)$ is:

$$\mathbf{u}(b^+) = V^T\,\mathbf{z}_1(b^+) + T^T(a^-)\,\mathbf{z}_2(b^+) \tag{86}$$

which must be matched to the toroidal flow in expression (79)

$$\mathbf{u}(b^+) = U^T(b^+)\,\mathbf{w}_1 \tag{87}$$

We thereby obtain the following simple system:

$$\mathbf{B}\,\mathbf{t} = \boldsymbol{\delta_t} \tag{88}$$

in which

$$\mathbf{t} = \begin{pmatrix} U^T(b^+) \\ T^T(a^-) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\delta_t} = V^T\,\mathbf{z}_1(b^+) \tag{89}$$

Each column of the 2×2 matrix $\mathbf{B}$ in (88) involves the corresponding basis vectors $\mathbf{w}_1$ and $-\mathbf{z}_2$ in (86-87). Expression (88) shows that the toroidal flow throughout the mantle is 'driven' by the surface toroidal flow V^T generated by rotating tectonic plates, which are themselves driven by the buoyancy-induced poloidal flow in the mantle. This surface coupling of poloidal and toroidal flow, due to the presence of surface tectonic plates, is discussed in the next section.

2.3.7 Incorporating tectonic plates as a surface boundary condition

The boundary conditions at the top of the mantle are more complex because of the presence of tectonic plates. If the plates are assumed to be an integral part of the underlying mantle flow and they fully participate in the flow, then a simple free-slip surface boundary condition would be appropriate. This free-slip assumption has often been employed in tomography-based flow studies (e.g., Hager *et al.*, 1985; Forte & Peltier, 1987; Hager & Richards, 1989; King & Masters, 1992). If, on the other hand, we recognise that the plates are mechanically and rheologically distinct from the underlying mantle, then a more realistic surface boundary condition that explicitly treats the effective rigidity of plates is required.

Essentially rigid tectonic plates are a fundamental aspect of convection dynamics in the Earth and, although their treatment lies beyond classical fluid mechanics theory, their dynamical impact on thermal convection in the mantle must be considered. One approach for modelling the coupling of convection and plates, developed by Ricard & Vigny (1989) and Gable *et al.* (1991), is based on an explicit treatment of the net vertical torque (or horizontal force) acting on the base of each surface plate. This method is a direct extension of the torque-balance analysis originally employed by Hager & O'Connell (1981) in their modeling of dynamic plate motions. The fundamental underlying assumption in the method developed by Ricard & Vigny is that the plate boundaries are completely stress-free. This method has been employed in a number of mantle-flow models over the past few years (e.g. Corrieu *et al.* 1994; Lithgow-Bertelloni & Richards 1998).

An alternative method for coupling the motions of rigid surface plates to buoyancy induced mantle flow was developed by Forte and Peltier (1991,1994). Only the main aspects of this method will be summarised here. It may be shown that for a given geometry of surface plates, the internal density anomalies $\rho_1(r,\theta,\phi)$ are partitioned into two families:

$$\rho_1(r,\theta,\phi) = \hat{\rho}_1(r,\theta,\phi) + \bar{\rho}_1(r,\theta,\phi), \tag{90}$$

where the density perturbations $\hat{\rho}_1(r,\theta,\phi)$ are obtained through a projection operator $\hat{P}$ as follows:

$$\hat{\rho}_1{}_s^t(r) = \hat{P}_{s\,t,\,l\,m}(r)\;\rho_1{}_l^m(r) \tag{91}$$

The calculation of the projection operator $\hat{P}_{s\,t,\,l\,m}(r)$ is given in *Forte and Peltier* (1994) and it depends on the geometry of the surface plates. The other component of the density anomalies $\bar{\rho}_1(r,\theta,\phi)$ is simply given by the expression:

$$\bar{\rho}_1{}_s^t(r) = \rho_1{}_s^t(r) - \hat{\rho}_1{}_s^t(r) \tag{92}$$

As Forte & Peltier (1994) show, the poloidal mantle flow field produced by the component $\hat{\rho}_1$ is consistent with the geometry of possible rigid plate motions at the surface whereas the one produced by $\bar{\rho}_1$ is orthogonal to any possible plate motion. In other words, the plates participate in the underlying flow driven by $\hat{\rho}_1$, while they resist the flow produced by $\bar{\rho}_1$. Hence, free-slip ($T^p(r_s) = 0$) and no-slip ($V^p(r_s) = 0$) surface boundary conditions are applied to model the internal flows driven by $\hat{\rho}_1$ and $\bar{\rho}_1$, respectively.

A simple application of the plate-projection operator is illustrated in Fig. 7, where we compare the buoyancy driven plate motions predicted on the basis of two hypothetical hot thermal anomalies under the Pacific plate which differ only in their spatial relationship to the nearby East Pacific Ridge (EPR). Both 'plumes' are constructed in an *ad hoc* manner as rectangular shaped, negative density anomalies with constant amplitude -0.1 g/cm^3 located in the depth range 400 to 750 km, equivalent to -2.6% relative density perturbation at 525 km depth. This amplitude is much greater than expected on the basis of the seismic tomography models (see Fig. 12 below) and is only used here for the purpose of illustration. The predicted mantle flow field varies linearly with the amplitude of the driving density anomaly, therefore a 'plume' anomaly which is ten times smaller (-0.01 g/cm^3) will produce surface motions ten times smaller than those shown in Fig. 7.

In the absence of plates, both of the hypothetical plumes (Figs. 7a,b) produce the same amplitude and pattern of horizontal flow divergence at the surface of the mantle (compare Figs. 7c,d). The situation changes radically when the present-day configuration of rigid plates is imposed at the top of the mantle. The plume which is offset from the nearest plate boundary (Fig. 7a), the EPR, drives a mantle flow field which produces only a small fraction of the horizontal surface divergence obtained in the absence of plates (compare Figs. 7c,e). The plate projection operator (91) maps most of the original plume density anomaly ρ_1 into the no-slip family of density anomalies $\bar{\rho}_1$ which cannot produce observable plate motions. According to the alternative plate-motion theory employed by Ricard & Vigny (1989), this hypothetical plume anomaly produces flow-induced driving torques on the overlying Pacific plate which nearly cancel because the centre of the upwelling is too far away from the nearest plate boundary. In contrast, the plume located under the EPR (Fig. 7b) produces horizontal divergence of the Pacific and Nazca plates with an amplitude comparable to that obtained in the absence of plates (compare Figs. 7d,f). In this case, a substantial fraction of the original density anomaly ρ_1 was mapped into the free-slip family of density anomalies $\hat{\rho}_1$. It should be noted that mutual interactions amongst the plates will generate non-zero far-field plate divergence and convergence (Fig. 7f, for example in the Indian Ocean) which is a consequence of the assumed rigidity of the plates.

From the perspective of the tectonic plates, the existence of a null-space corresponding to the family of internal density anomalies $\hat{\rho}_1$ (92) which cannot drive observable motions (as in Fig. 7e) constitutes a fundamental non-uniqueness in the interpretation of plate tectonics. The present-day plate motions (DeMets et al. 1990), or the geologic reconstructions of the history of plate motions (Gordon & Jurdy 1986), therefore provide completely non-unique constraints on the mantle density anomalies generated by the thermal convection process. Other convection-related observables (e.g. global gravity anomalies) are required to constrain the density anomalies which fall into the null-space $\hat{\rho}_1$. A futher exploration of the implications of the non-unique interpretation of plate motions is presented by Forte & Peltier (1994).

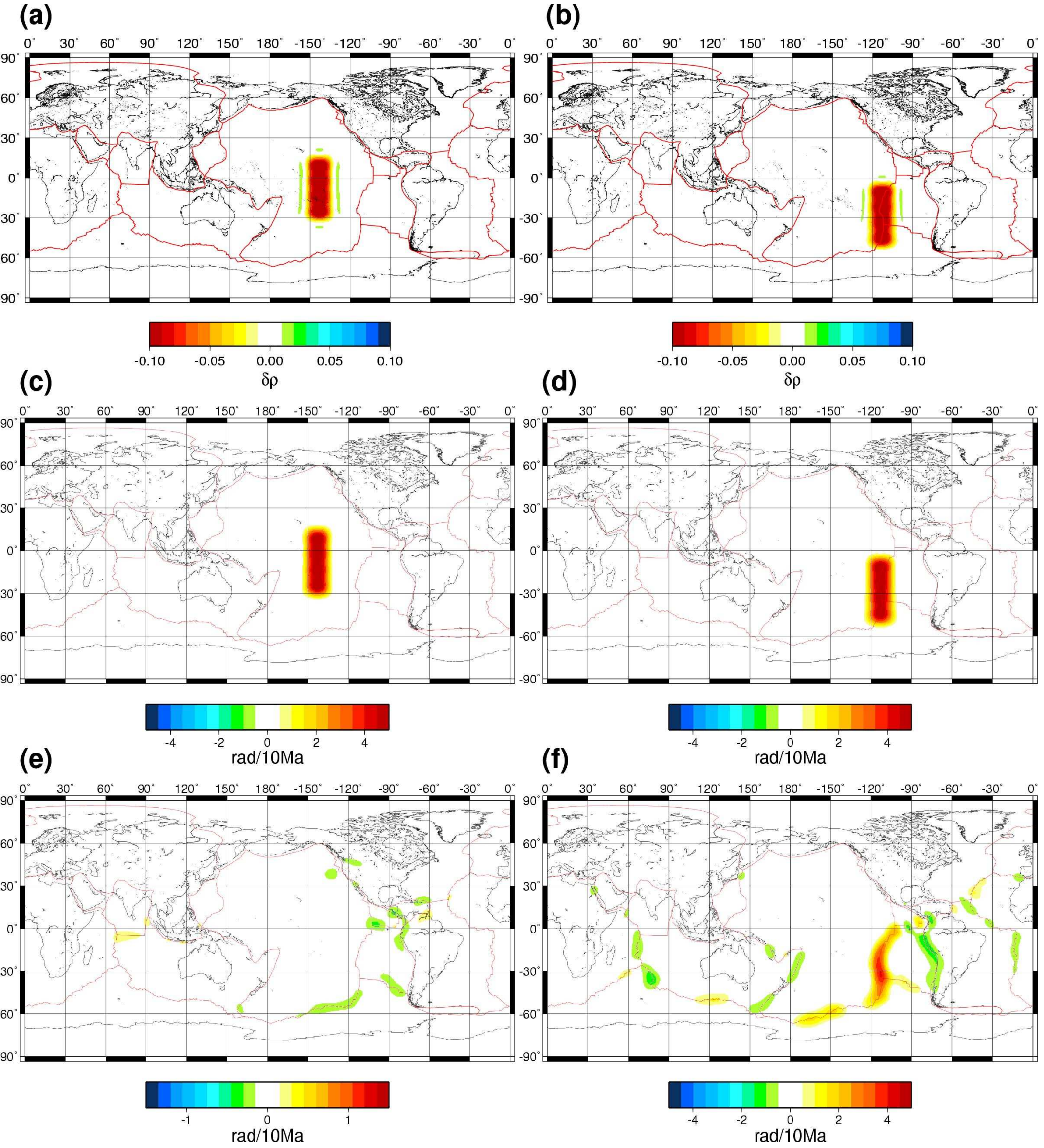

Figure 7: Interaction of hypothetical upwelling plumes with rigid surface plates. Maps (a) and (b) show two rectangular shaped, negative density anomalies (the hypothetical 'plumes') with amplitude -0.1 g/cm^3 assumed to be constant in the depth interval 400 to 750 km. Both plume anomalies have been expanded in terms of spherical harmonics up to degree and order 32. Maps (c) and (d) show the corresponding patterns of predicted horizontal surface divergence for a hypothetical mantle with a free-slip surface boundary (no plates) where the predicted motions are calculated up to harmonic degree 32. The mantle flow calculation was carried out assuming a depth-dependent viscosity variation shown below in Fig. 9. Maps (e) and (f) show the corresponding patterns of plate divergence (up to harmonic degree 32) when the present-day configuration of rigid tectonic plates is imposed as a surface boundary condition in the mantle flow calculation.

The plate motions which are driven by the buoyancy induced (poloidal) mantle flow will necessarily posses a toroidal component. This toroidal component is directly coupled to the poloidal component of the plate motions as follows:

$$(\hat{\boldsymbol{r}} \cdot \boldsymbol{\nabla} \times \mathbf{v})_s^t = C_{s\,t,\,l\,m}\ (\boldsymbol{\nabla}_H \cdot \mathbf{v})_l^m \ , \tag{93}$$

in which the spherical harmonic coefficients of the radial vorticity of the plate velocity field **v** are linearly dependent on the harmonic coefficients of the horizontal divergence of the plate motions. The vorticity and divergence fields are themselves directly dependent on the toroidal and poloidal scalar representations, respectively, of the plate velocity field (Forte & Peltier 1987). It is important to note that expression (93) is only valid for toroidal harmonic degree $s \geq 2$ and that the coupling matrix $C_{s\,t,\,l\,m}$ is dependent only on the geometry of the plates (Forte & Peltier 1994).

As discussed above in section (2.3.6), the viscous coupling between the plates and underlying mantle will result in the downward propagation of the surface toroidal flow into the mantle. The surface boundary condition (74) requires that the toroidal mantle flow match the toroidal component of plate motions, which is equivalent to their radial vorticity (93):

$$V^T = U^T(a^-) = \frac{\imath\, r_a}{\Omega_1}\,(\hat{\boldsymbol{r}} \cdot \boldsymbol{\nabla} \times \mathbf{v})_l^m \ , \tag{94}$$

where $\imath = \sqrt{-1}$ and $r_a = a$ is the mean radius of the solid surface (6368 km).

The Green function solutions of the poloidal and toroidal flow equations (see section 2.3.6), may be used to describe the viscous response of the mantle to an arbitrary distribution of internal density anomalies ρ_1 as follows:

$$\begin{pmatrix} U_\ell^{\alpha\, m}(r) \\ T_\ell^{\alpha\beta\, m}(r) \end{pmatrix} = \int_b^a \begin{pmatrix} \mathsf{U}_\ell^{\alpha}(r, r') \\ \mathsf{T}_\ell^{\alpha\beta}(r, r') \end{pmatrix} \, {\rho_1}_\ell^m(r')\, \mathrm{d}r' \ , \tag{95}$$

where $U_\ell^{\alpha\, m}(r)$ and $T_\ell^{\alpha\beta\, m}(r)$ are the generalised spherical harmonic coefficients of the velocity and stress fields (31, 32), respectively, and $\mathsf{U}_\ell^{\alpha}(r, r')$ and $\mathsf{T}_\ell^{\alpha\beta}(r, r')$ are the corresponding (poloidal) Green functions.

Tectonic plates may well represent the most extreme manifestation of lateral variations in rheology in the convecting mantle. The plate coupling procedure described above thus provides an explicit means of incorporating the dynamical impact of such extreme lateral heterogeneity in the flow calculations which will be presented below. It is nonetheless true that the equations (35,36) governing flow in the mantle do not include lateral variations in viscosity and hence buoyancy forces will only directly excite a poloidal flow. In these calculations, toroidal mantle flow is excited passively from above, via the rotating surface plates.

A more general flow theory which models the effects of lateral viscosity variations throughout the mantle (e.g., Martinec *et al.* 1993; Forte & Peltier 1994) is required to describe the direct excitation of toroidal flow by mantle buoyancy forces. The dynamical implications of large-amplitude lateral viscosity variations throughout the mantle have been studied by Moucha *et al.* (2007). Despite the considerably more complex theory (relative to equations 35,36) needed to calculate flow in a 3-D viscosity field, Moucha *et al.* find the impact on predicted surface observables such as topography and geoid are relatively small. This finding provides important support for the modelling and interpretation of geodynamic observables using the plate-coupled flow theory with purely depth dependent viscosity in the mantle. The results obtained on the basis of this simplified treatment of mantle dynamics will be presented in the sections which follow.

2.4 Geodynamic response functions for the mantle

The theoretical relationship between the mantle density anomalies and the principal convection related surface observables (i.e. geoid or gravity anomalies, dynamic surface and CMB topography, plate motions) may be summarised in terms of geodynamic response or kernel functions . (See Hager & Clayton 1989, and Forte 2000, for reviews.) These kernel functions are calculated in the spherical harmonic spectral domain using the viscous flow Green functions defined in expression (95). The principal assumption we shall make in calculating these geodynamic kernels - in addition to the assumption of purely depth dependent viscosity below the plates - is that the 670 km seismic discontinuity is not a barrier to radial mantle flow. This whole-mantle flow assumption has been recently subjected to a series of direct tests based on joint inversions of global seismic and geodynamic data sets (Simmons *et al.* 2006). It is found that the most satisfactory reconciliation of the global seismic and geodynamic constraints on 3-D mantle structure is obtained with whole-mantle flow and significantly poorer fits are obtained with flow models which assume vertical flow barriers at 670 km depth or within the lower mantle (e.g. at 1800 km depth).

The geoid kernels $G_\ell(\eta; r)$ relate the spherical harmonic coefficients of the nonhydrostatic geoid field, δN_ℓ^m, to the spherical harmonic coefficients of the density perturbations $\delta\rho_\ell^m(r)$ as follows:

$$\delta N_\ell^m = \frac{3}{(2\ell+1)\bar{\rho}} \int_{r_{\rm cmb}}^{r_s} G_\ell(\eta; r)\, \delta\rho_\ell^m(r)\, {\rm d}r \;, \tag{96}$$

where $\bar{\rho} = 5.515\ \mathrm{Mg/m^3}$ is the mean density of the Earth, r_s is the mean radius of the solid surface, and $r_{\rm cmb}$ is the mean radius of the CMB. The geoid kernels are a functional of the non-dimensional radial mantle viscosity profile $\eta(r)/\eta_o$. The free-air gravity anomaly coefficients δG_ℓ^m are related to the nonhydrostatic geoid coefficients δN_ℓ^m as follows:

$$\delta G_\ell^m = (\ell - 1)\frac{g_o}{R_o}\, \delta N_\ell^m \;, \tag{97}$$

where $g_0 = 9.82\ \mathrm{m/s^2}$ (982,000 mGal) is the mean gravitational acceleration at Earth's surface and $R_0 = 6371$ km is Earth's mean radius. Through the combination of equations (96) and (97), we can relate the gravity anomaly coefficients to the internal density perturbations.

As discussed in the preceding section, a dynamical flow model which incorporates rigid plate motions must include mixed free-slip and no-slip surface boundary conditions. In the calculation of the flow-related observables, we must therefore determine the separate contributions provided by the density anomalies $\delta\hat{\rho}$ and $\delta\bar{\rho}$ which are convolved with kernel functions derived with free-slip and no-slip surface boundaries, respectively.

For example, in the case of the nonhydrostatic geoid anomalies, we rewrite equation (96) as follows:

$$\delta N_\ell^m = \frac{3}{(2\ell+1)\bar{\rho}} \int_{r_{\rm cmb}}^{r_s} \left[\hat{G}_\ell(\eta; r)\, \delta\hat{\rho}_\ell^m(r) + \bar{G}_\ell(\eta; r)\, \delta\bar{\rho}_\ell^m(r) \right] {\rm d}r \,, \tag{98}$$

where the geoid kernels $\hat{G}_\ell(\eta; r)$ and $\bar{G}_\ell(\eta; r)$ are calculated with free-slip and no-slip surface boundaries, respectively. As shown in Forte & Peltier (1994), the density anomalies $\delta\hat{\rho}$ are spatially correlated with the positions of divergent or convergent plate boundaries. Subducting slabs and mid-ocean ridge heterogeneity are therefore resolved as part of the $\delta\hat{\rho}$ anomalies.

The surface topography kernel functions $T_\ell(\eta; r)$ relate the spherical harmonic coefficients of the flow-induced surface topography δa_ℓ^m to the spherical harmonic coefficients of the internal density anomalies as follows:

$$\delta a_\ell^m = \frac{1}{\Delta\rho_{mo}} \int_{r_{\rm cmb}}^{r_s} \left[\hat{T}_\ell(\eta; r)\, \delta\hat{\rho}_\ell^m(r) + \bar{T}_\ell(\eta; r)\, \delta\bar{\rho}_\ell^m(r) \right] {\rm d}r \,, \tag{99}$$

where $\Delta\rho_{mo} = 2.2\,{\rm Mg/m^3}$ is the density jump across the mantle-ocean boundary and $\hat{T}_\ell(\eta; r)$ and $\bar{T}_\ell(\eta; r)$ are calculated with free-slip and no-slip surface boundaries, respectively.

Similarly, the spherical harmonic coefficients of the flow-induced CMB topography δb_ℓ^m may be expressed in terms of CMB topography kernels $B_\ell(\eta; r)$ as follows:

$$\delta b_\ell^m = \frac{1}{\Delta\rho_{cm}} \int_{r_{\rm cmb}}^{r_s} \left[\hat{B}_\ell(\eta; r)\, \delta\hat{\rho}_\ell^m(r) + \bar{B}_\ell(\eta; r)\, \delta\bar{\rho}_\ell^m(r) \right] {\rm d}r \,, \tag{100}$$

where $\hat{B}_\ell(\eta; r)$ and $\bar{B}_\ell(\eta; r)$ are calculated with free-slip and no-slip surface boundaries, respectively, and $\Delta\rho_{cm} = -4.43\,{\rm Mg/m^3}$ is the density jump across the CMB. As in the case of the nonhydrostatic geoid in equation (98), the dependence on viscosity of the predicted surface and CMB topography appears explicitly through the non-dimensional profile $\eta_(r)/\eta_o$.

Surface plate velocities $\mathbf{v}$ may be characterised by their horizontal divergence field $\boldsymbol{\nabla}_H \cdot \mathbf{v}$ (Forte & Peltier 1987). The horizontal divergence kernels $D_\ell(\eta; r)$ relate the spherical harmonic coefficients of the predicted plate divergence $(\boldsymbol{\nabla}_H \cdot \mathbf{v})_\ell^m$ to the spherical harmonic coefficients of the density perturbations as follows:

$$(\boldsymbol{\nabla}_H \cdot \mathbf{v})_\ell^m = \frac{g_o}{\eta_o} \int_{r_{\rm cmb}}^{r_s} D_\ell(\eta; r)\, \delta\hat{\rho}_\ell^m(r)\, {\rm d}r \,. \tag{101}$$

in which the free-slip component of the mantle density anomalies $\delta\hat{\rho}$, defined above in expression (91), is required to model the surface divergence of the plates. We also note the presence of the reference scaling value of viscosity η_0. This presence implies that plate motions, unlike the other observables in equations (96-100), will be dependent on the absolute value of mantle viscosity.

2.5 Depth dependence of mantle viscosity

Understanding the long time scale rheology of the mantle as represented by its effective viscosity, is a central and enduring problem in global geophysics. The diverse methods and data sets which have been employed to constrain mantle viscosity have been a source of ongoing contention and debate. The importance and intensity of this debate are an apt reflection of the fundamental role of mantle viscosity in controlling a wide array of geodynamic processes. For example, millennial time scale glacial isostatic adjustment (GIA) processes such as Pleistocene and Holocene sea-level variations and related anomalies in Earth's gravitational field and rotational state are known to be strongly dependent on the depth dependence of mantle viscosity. On much longer, million to hundred-million year time scales, viscosity exerts fundamental control on the dynamics of mantle convection and on the corresponding evolution of the thermal and chemical state of Earth's interior. The very long time scale implications of mantle viscosity also include fundamental surface geological and geophysical processes, such as global scale epeirogeny and associated sea level changes, global geoid anomalies and tectonic plate motions.

The spatial variation of viscosity in Earth's mantle can, in principle, be determined from a consideration of the microphysical properties which control the steady state creep of mantle rocks (e.g. Nicolas & Poirier 1976; Sammis et al. 1977). In practise, however, the microphysical models of mantle viscosity depend on knowledge of a number of critical physical properties, such as grain size and activation volume and energy, which are poorly known in the mantle and hence require extrapolations of laboratory creep experiments which greatly exceed the physical conditions (i.e. pressure, temperature, or creep rate) under which these experiments were originally performed. It is nonetheless possible to show, on the basis of microphysical considerations, that the global horizontal average viscosity may increase by up to three orders of magnitude across the mantle (e.g. Ranalli 2001). More precise determinations of the radial viscosity profile require additional constraints from geophysical data sets which are sensitive to the long term rheology of the mantle.

The first and most influential geophysical contribution to our understanding of mantle viscosity is Haskell's (1935) study of Fennoscandian post-glacial uplift which was found to require an average viscosity of 10^{21} Pa s down to a depth approximately equal to the horizontal dimension of the surface load (about 1000 to 1500 km). McConnell's (1968) spectral analysis of the Fennoscandian uplift data led to the inference that mantle viscosity increased significantly with depth while at the same time still satisfying the Haskell average. The Haskell average value of viscosity over the upper 1000 to 1200 km of the mantle was also supported in subsequent work by O'Connell (1971). These initial studies which supported the compatibility between the Haskell constraint on viscosity and significant increases of viscosity in the lower mantle were displaced by subsequent analyses by Cathles (1975) and Peltier & Andrews (1976) who concluded that viscosity was nearly constant from the base of the lithosphere to the CMB. This view of mantle viscosity characterised by only moderate increases with depth was reinforced in later studies involving larger GIA data sets (e.g. Wu & Peltier 1983; Tushingham & Peltier 1991). A review of the GIA constraints on viscosity may be found in Mitrovica (1996).

The first inferences of the depth variation of mantle viscosity based on surface data sets associated with mantle convection (e.g. Hager 1984; Richards & Hager 1984; Ricard et al. 1984) disagreed strongly with the GIA inferences of mantle viscosity which prevailed since the earliest studies by Cathles, Peltier and collaborators. These convection-based analyses of viscosity, in particular the first study of Earth's global geoid anomalies by Hager (1984), suggested a large increase of viscosity in the sub-lithospheric mantle with an approximately factor of 30 increase in the average viscosity of the lower mantle relative to the upper mantle. This conclusion was supported by subsequent studies by Forte & Peltier (1987) and Forte et al. (1991) who extended the viscous flow modelling to include the plate velocities. The geodynamic inferences for significant (one to two order of magnitude) increases in the average viscosity of the lower mantle relative to that of the upper mantle were reinforced by subsequent analyses carried out by Ricard et al. (1989), Ricard & Vigny (1989), Hager & Clayton (1989), Forte & Peltier (1991), Forte et al. (1993a,1994), King & Masters (1992), Corrieu et al. (1994), Thoraval & Richards (1997). A survey of some of the earliest viscosity inferences derived on the basis of global geoid anomalies are shown in Fig. 8. Reviews of mantle viscosity inferences based on either convective flow or GIA data sets may by found in King (1995) and Kaufmann & Lambeck (2000).

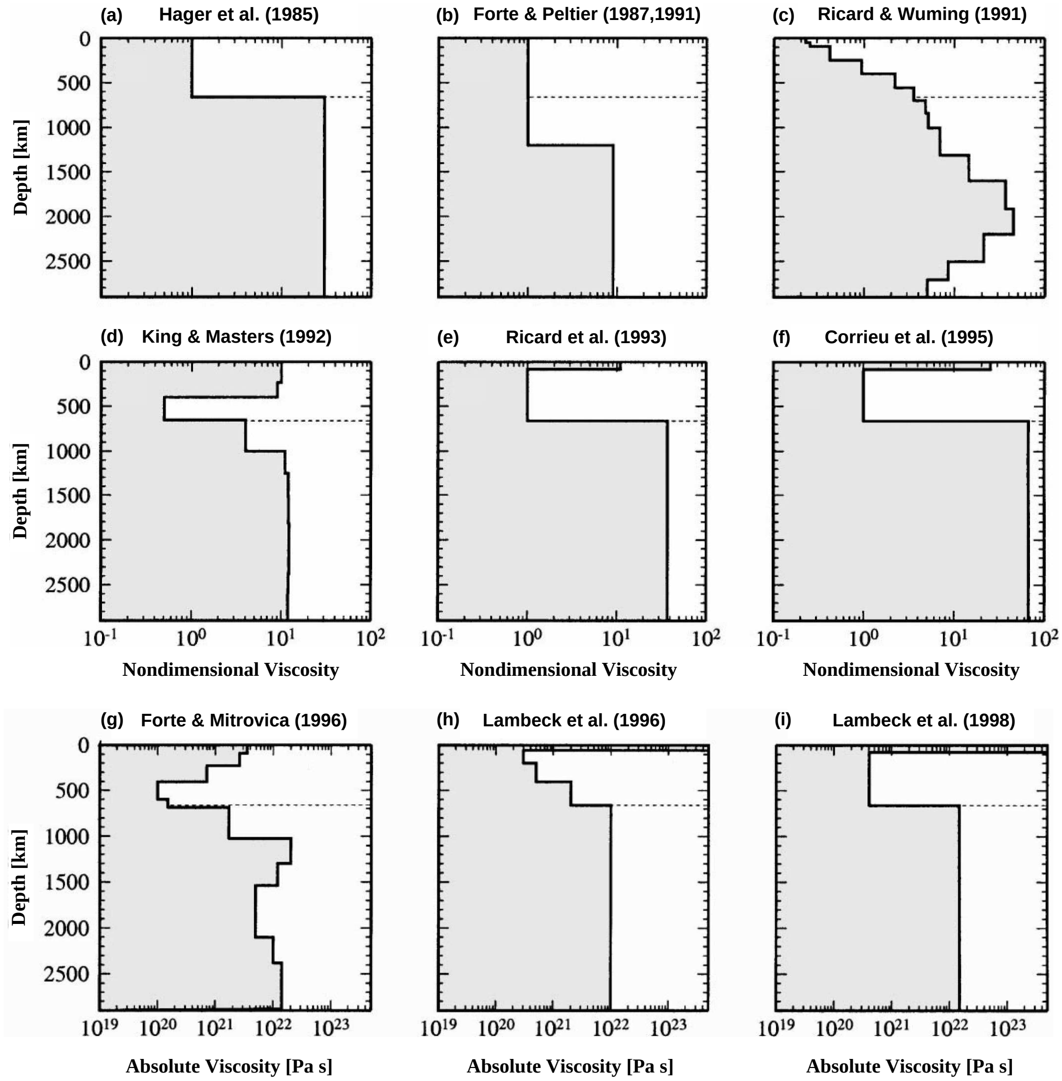

Figure 8: A survey of previous geodynamic inferences of depth dependent mantle viscosity. Frames (a) to (f) show relative (nondimensional) variations of viscosity with depth derived on the basis of global long wavelength geoid anomalies. Frames (g) to (i) show radial (absolute) viscosity profiles inferred on the basis of GIA data (in addition to geoid data in the case of frame g). The GIA data sets constrain the absolute value of viscosity whereas the steady-state viscous flow modelling of the geoid data only constrain relative variations of viscosity with depth. (Figure adapted from Kaufmann & Lambeck, 2000)

Additional geodynamic considerations which suggest a substantial increase of viscosity with depth are based on a wide variety of analyses which include the stability of the hot spot reference frame (e.g. Richards 1991; Quéré & Forte 2006), long-term rates of polar wander (e.g. Sabadini & Yuen 1989; Spada et al. 1992; Richards et al. 1997; Steinberger & O'Connell 1997), and the planform of mantle convection (e.g. Zhang & Yuen 1995; Bunge et al. 1996; Forte & Mitrovica 2001).

The inference of mantle viscosity from convection-related surface data, considered in the context of tomography-based mantle flow models, was initially carried out through an effectively trail-and-error fitting procedure (e.g., Forte & Peltier 1987,1991; Hager & Clayton 1989; Forte *et al.* 1993). Subsequently, a number of formal mathematical inversions of the geodynamic data were carried out (e.g., Forte *et al.* 1991; Ricard & Wuming 1991; King & Masters 1992; Corrieu *et al.* 1994; King 1995; Forte & Mitrovica 1996; Kido *et al.* 1998; Forte 2000; Panasyuk & Hager 2000). The viscosity profiles obtained by inverting geodynamic data sets (Fig. 8) are all characterised by strong overall increases in viscosity with depth, but it is also clear that there are significant differences in the detailed character of these viscosity inferences. There are many factors which can contribute to these differences and they include, most notably, the use of: different tomography-based inferences of mantle density anomalies; different combinations of geodynamic data (e.g., geoid only, geoid plus plate motions, geoid plus dynamic topography, all of the above); different inversion methods (e.g., Monte-Carlo, Bayesian, genetic, Occam algorithms). The formulation of the viscous flow models themselves can also play a critical role in explaining these differences, since it is known that the use of different surface boundary conditions (e.g. free-slip, no-slip, freely rotating tectonic plates) will lead to rather different viscosity inferences (e.g. Thoraval & Richards 1997). Finally, the limited resolving power of the geodynamic constraints precludes a direct interpretation of viscosity at any given depth in the mantle. For example, Forte & Mitrovica (2001) found that the combined constraints provided by global free-air gravity anomalies, tectonic plate motions, and excess ellipticity of the CMB could only resolve average values of viscosity in rather broad depth intervals.

The discordant viscosity inferences derived from either GIA or convection-related data sets were initially regarded as evidence for the transient nature of mantle viscosity over the time scales that separate these two processes (e.g. Sabadini et al. 1985; Peltier 1985). The necessity for time-dependent viscosity in the mantle was later weakened by two independent analyses of the GIA data sets. The first argument against transient viscosity was provided by an analysis of differential late Holocene sea-level high-stands in the Australian region by Nakada and Lambeck (1989), who showed that these GIA data required a nearly two order of magnitude increase of viscosity across the mantle. The second independent argument was provided by Mitrovica (1996), who showed that previous inferences of a nearly constant-viscosity mantle (e.g. Wu & Peltier 1983; Tushingham & Peltier 1991) were conditioned by a misinterpretation of the original Haskell (1935) constraint. The latter was erroneously assumed to apply only to the top 670 km of the mantle (i.e. the upper mantle), rather than to the average value of viscosity down to depths of about 1000 to 1500 km (Mitrovica 1996).

The recognition that GIA data are not incompatible with large increases of mantle viscosity with depth spurred new efforts to reconcile GIA and convection data sets with a single profile of mantle viscosity. The first efforts were undertaken by Forte & Mitrovica (1996) and Mitrovica & Forte (1997). Despite the very different time and spatial scales over which these processes operate, these authors found that it was possible to simultaneously explain both decay times associated with the post-glacial uplift of Hudson Bay and Fennoscandia, and long-wavelength free-air gravity anomalies associated with mantle convection with a single viscosity profile characterised by an approximately two order of magnitude increase with depth which also satisfied the original Haskell constraint on mantle viscosity. These joint GIA-convection inversions culminated in the recent study by Mitrovica & Forte (2004) in which the range of viscosity profiles consistent with both families of surface data are summarised below in Fig. reffig-visc. The convection data employed in these viscosity inversions are shown in Fig. 2. The GIA data include the Fennoscandian relaxation spectrum (FRS) and a set of decay times determined from the postglacial sea-level history in Hudson Bay and Sweden.

The inverted viscosity profiles are characterised by a three orders of magnitude increase from the upper mantle (mean value of $\sim 4 \times 10^{20}$ Pa s) to a high-viscosity ($> 10^{23}$ Pa s) peak at 2000 km depth. Below 2000 km, the viscosity shows a significant, two to three orders of magnitude, reduction toward the CMB. Similar radial variations in lower-mantle viscosity have been inferred in other independent studies (e.g., Ricard & Wuming 1991). The preferred viscosity profile (solid black line, Fig. 9) provides an optimal fit to both (GIA and convection) families of data and, unless stated otherwise, this profile will be employed to calculate all convection related related surface observables presented below.

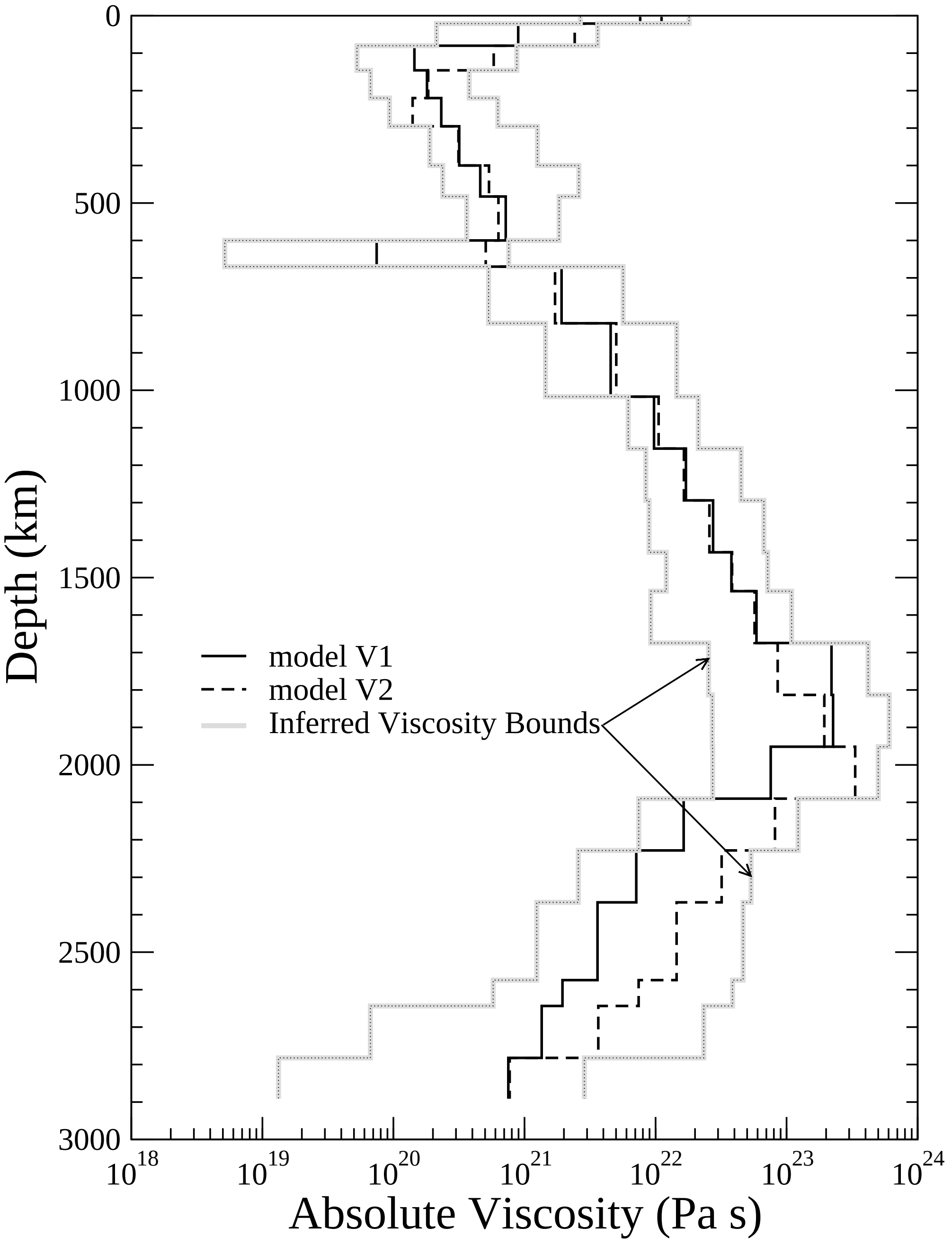

Figure 9: Results of inversions of the GIA and convection data sets. Full details of the inversions are presented in Mitrovica & Forte (2004). The convection data are interpreted in terms of a tomography-based mantle flow model in which the density anomalies are derived from the Grand (2002) tomography model (Fig. 3b). The thick gray lines illustrate the range of allowable values of mantle viscosity which are consistent with the joint GIA-convection constraints. The solid black line (model V1) is the preferred viscosity profile, on the basis of the Grand (2002) tomography model, which provides an optimal fit to the entire suite of geodynamic data. The dashed line (model V2) was inverted with strong smoothing constraints applied across the horizon separating upper and lower mantle and yields slightly lower fits to the combined GIA-convection data sets (Forte *et al.* 2010).

The geodynamic response or kernel functions calculated on the basis of the preferred mantle viscosity profile (model 'V1' in Fig. 9) are illustrated below in Fig. 10. Comparing the free-slip and no-slip geoid kernels, we note that the largest contribution from density anomalies in the top half of the mantle is provided by those anomalies $\delta\hat{\rho}$ which are efficient in driving observable plate motions. These anomalies correspond to subducting slabs and upper-mantle plumes below the mid-ocean ridges. In the bottom half of the mantle the density anomalies belonging to the $\delta\bar{\rho}$ family provide the largest contribution to the surface geoid or gravity anomalies. These $\delta\bar{\rho}$ anomalies produce no observable plate motions. At sufficiently short horizontal wavelengths (corresponding to harmonic degrees $\ell \geq 32$) we note that the distinction between $\delta\hat{\rho}$ and $\delta\bar{\rho}$ anomalies begins to disappear: both provide equal contributions to the surface geoid or gravity anomalies. The free-slip and no-slip surface topography kernels show that the density anomalies $\delta\bar{\rho}$ provide the strongest contribution to dynamic surface topography, especially at long horizontal wavelengths (corresponding to the degree range $\ell = 2 - 8$). At short horizontal wavelengths, the distinction between the two classes of density anomalies again ceases to be important.

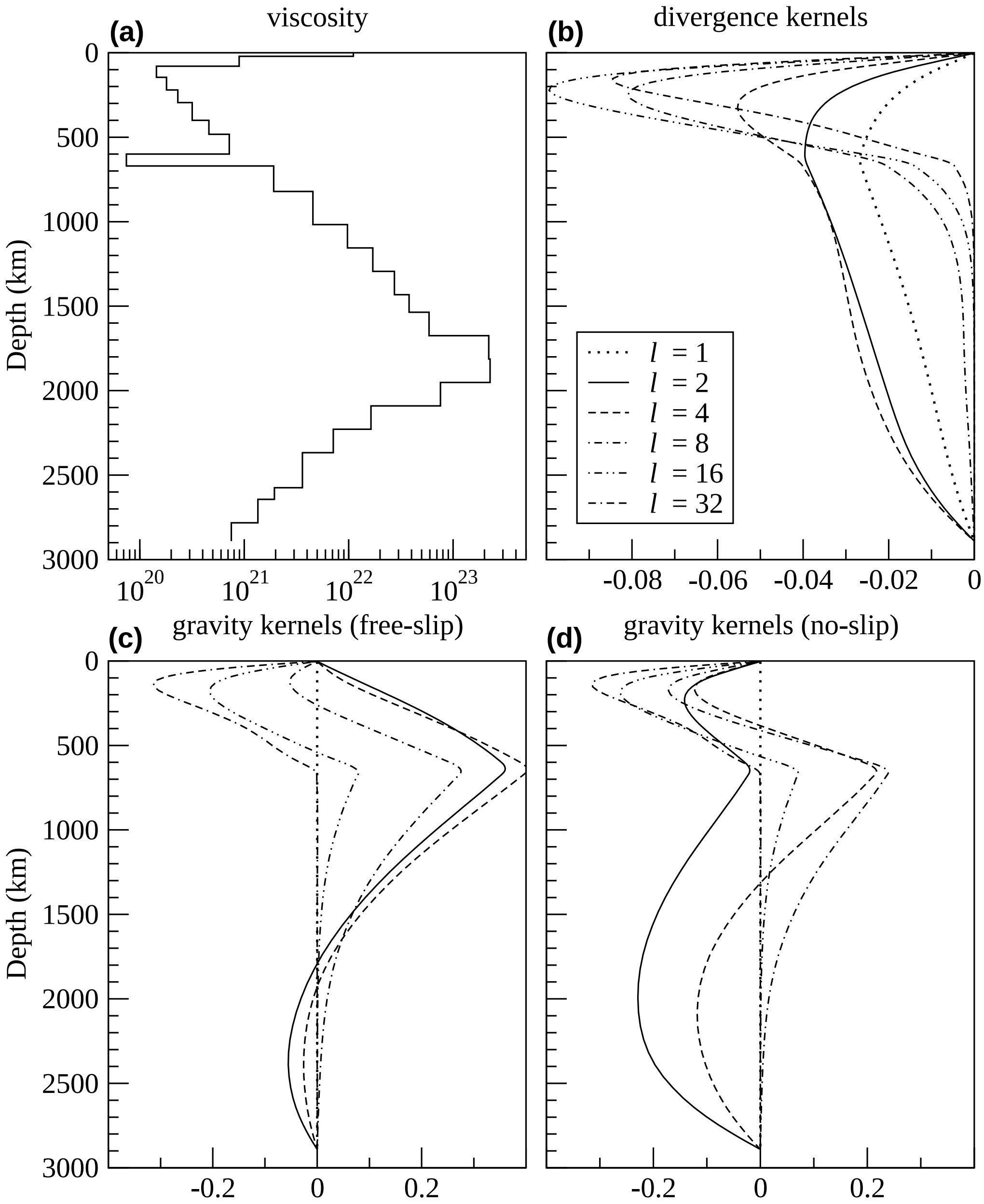

Figure 10: Geodynamic kernel functions. The geodynamic response functions shown here are calculated on the basis of the preferred GIA-convection-inferred viscosity profile shown in frame **a** (same as model 'V1' in Fig. 9). The kernels corresponding to harmonic degrees $\ell = 1, 2, 4, 8, 16, 32$ are identified by the legend in frame b (repeated below in frame e). The horizontal divergence kernels, defined in (101), are shown in frame **b**. The gravity kernels, defined in (98), for free-slip and no-slip surface boundary conditions are shown in frames **c** and **d**, respectively.

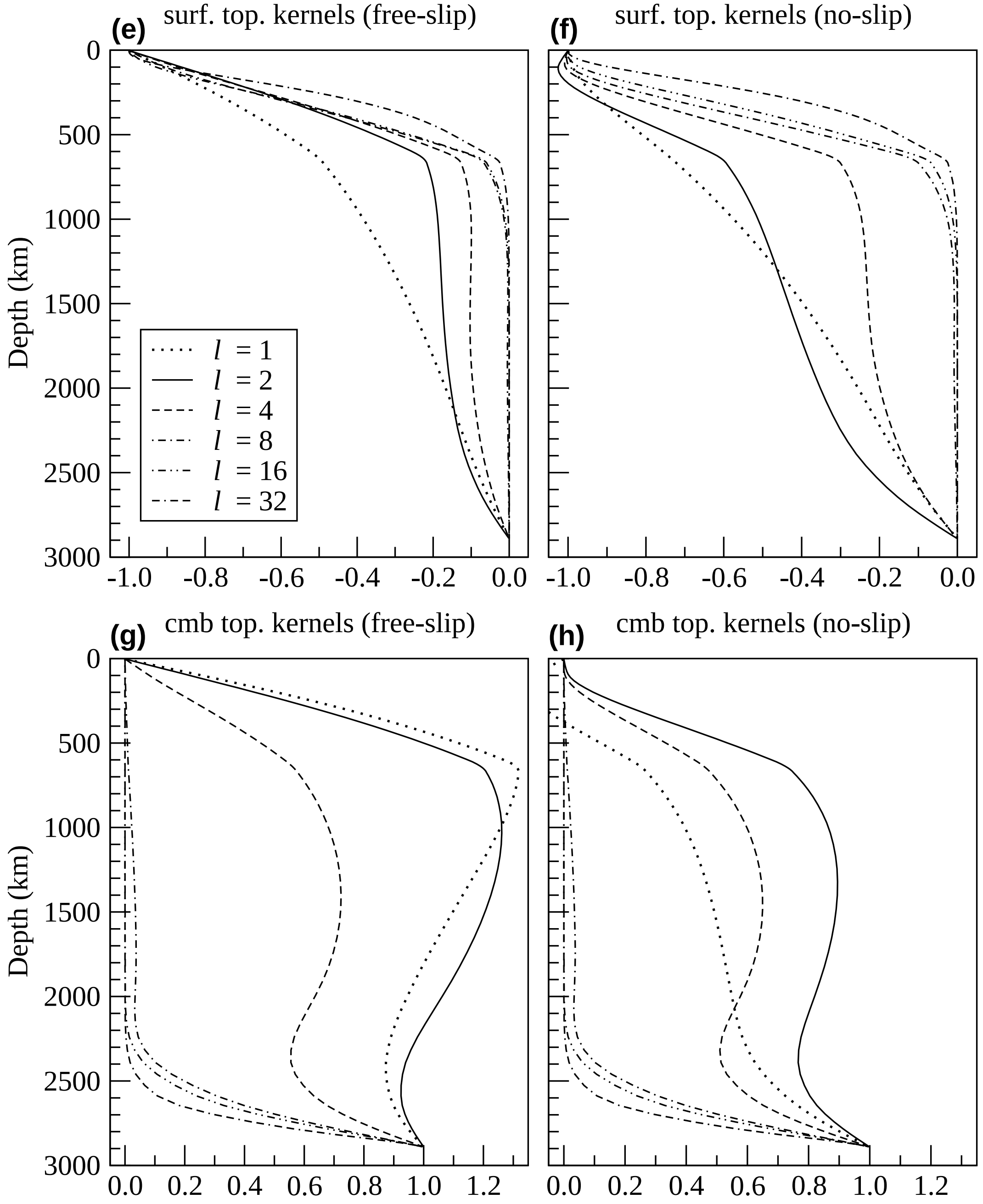

Figure 10: - *continued* - The surface topography kernels, defined in (99), for free-slip and no-slip conditions are shown in frames **e** and **f**, respectively. The kernels for the dynamic CMB topography, defined in (100) are shown in frame **g** and **h** for free-slip and no-slip conditions respectively.

3 Modelling geodynamic observables with seismic tomography

We will now consider the extent to which the 3-D mantle structure resolved in recent seismic tomographic models may be used to explain the global convection-related data sets shown in Fig. 2. The geodynamic response functions obtained on the basis of the preferred radial viscosity profile (Fig. 10) will be used to establish the connection between the seismically inferred heterogeneity in the mantle and the surface geodynamic observables. A broad selection of seismic shear-velocity models, obtained in independent studies using diverse and complementary seismic data and using different seismic modelling and inversion techniques, will be employed here. Seismic shear velocity heterogeneity is the initial focus of this tomography-based modelling because it is expected that S-wave anomalies δV_S will be most sensitive to the temperature anomalies maintained by the mantle convection process (e.g., Röhm *et al.* 2000; Forte & Perry 2000; Forte *et al.* 2002).

3.1 Seismic heterogeneity models

The S-wave tomography models employed here are, in chronological order: model *S20A* (isotropic version) from Ekström & Dziewonski (1998), model *S20RTS* from Ritsema *et al.* (1999), model *SAW24* from Mégnin & Romanowicz (2000), model *SB4_L18* from Masters *et al.* (2000), model *TX2002* from Grand (2002), and model *J362D28* from Antolik *et al.* (2003). The spatial variation of seismic heterogeneity in these S-wave models is parametrised in terms of a wide assortment of basis functions, ranging from B-splines to piecewise-discontinuous layers in the radial direction and spherical harmonics to equal-area cells in the horizontal direction.

To facilitate a direct comparison amongst these seismic heterogeneity models and the subsequent usage of the models in viscous flow calculations, all models have been projected onto a common set of radial and horizontal basis functions as follows:

$$\frac{\delta V_S}{V_S} = \sum_{k=0}^{32} \sum_{\ell=1}^{20} \sum_{m=-\ell}^{+\ell} {}_k c_\ell^m \, T_k(x) \, Y_\ell^m(\theta, \varphi) \tag{102}$$

in which $T_k(x)$ are Chebyshev polynomials in the normalised radius $x = (2r - r_{surf} - r_{cmb})/(r_{surf} - r_{cmb})$ and Y_ℓ^m are ordinary spherical harmonics. Notice that all models are expanded up to order 32 in the radial Chebyshev polynomials and up to degree 20 in the surface spherical harmonics.

When each of the S-wave heterogeneity models are projected onto the representation (102), the depth variations of the root-mean-square (rms) amplitudes of the velocity anomalies are shown in Fig. 11. Here we note that, with the exception of the top 300 km in the mantle, the average amplitudes of the relative perturbations of seismic shear velocity $\delta V_S/V_S$ show significant differences in the mantle, especially in the transition zone region (400 - 1000 km depth) and at the base of the mantle (2500 - 2891 km depth).

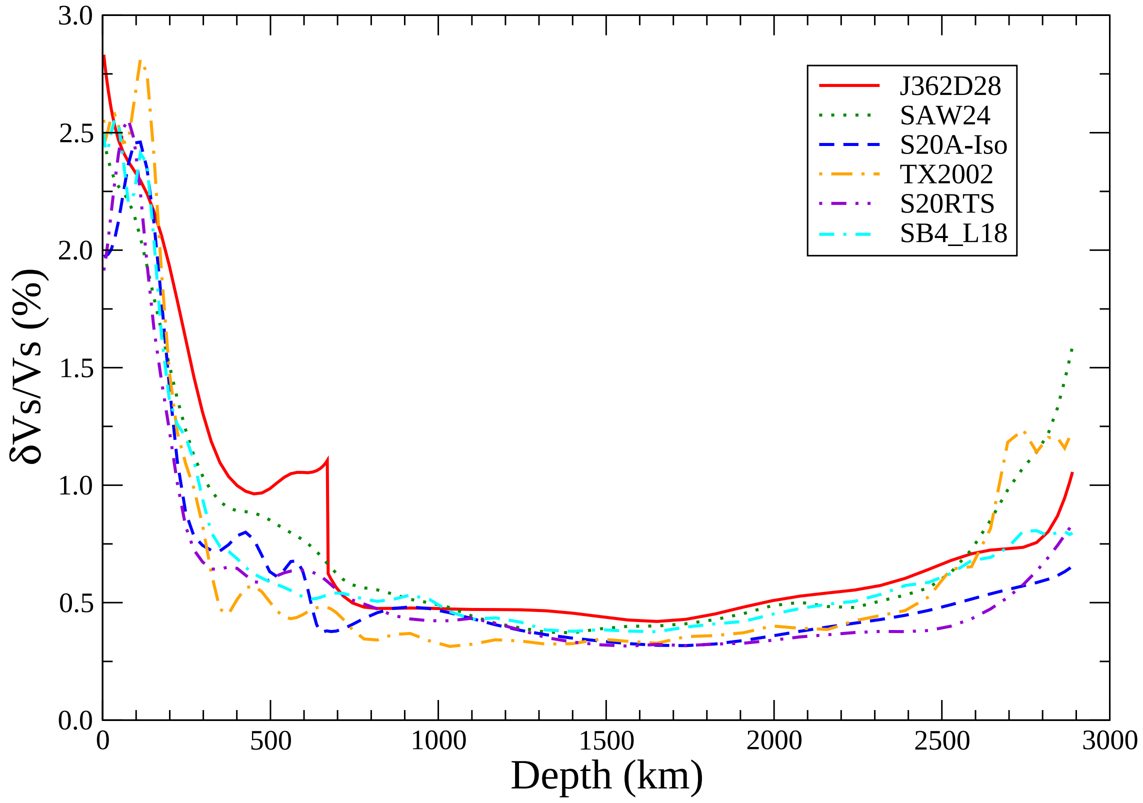

Figure 11: Amplitude of S-wave heterogeneity in the mantle. Plotted here is the root-mean-square (rms) amplitude of relative perturbations of seismic shear velocity obtained from the following tomography models: *J362D28* (Antolik *et al.* 2003), *SAW24* (Mégnin & Romanowicz 2000), *S20A-Iso* (Ekström & Dziewonski 1998), *TX2002* (Grand 2002), *S20RTS* (Ritsema *et al.* 1999), *SB4_L18* (Masters *et al.* 2000). All tomography models have been projected onto the spatial parametrisation given by expression (102).

3.2 Mantle density anomalies

The mantle density anomalies required in modelling the convection-related surface observables may be derived from the seismic tomography models using experimental and theoretical results from mineral physics (e.g., Karato 1993; Čadek *et al.* 1994; Wang & Weidner 1996; Sobolev *et al.* 1997; Stacey 1998; Zhang & Weidner 1999; Trampert *et al.* 2001; Stixrude & Lithgow-Bertelloni 2001; Jackson 2001; Karato & Karki 2001; Oganov *et al.* 2001a,b). This derivation would, however, require that the amplitudes of seismic velocity anomalies are well constrained and that all relevant mineralogical variables in the mantle (e.g., reference composition, temperature and equation of state parameters) are sufficiently well known. To satisfy all these requirements at the same time is a great challenge, but efforts to simultaneously interpret mantle density and seismic velocity anomalies using mineral physics data are showing much promise (e.g., Forte *et al.* 1994, Forte & Woodward 1997b, Simmons *et al.* 2007, 2009, 2010).

An alternative method for inferring mantle density anomalies is based on the direct inversion of the convection-related geophysical observables (e.g. Forte *et al.* 1993; Corrieu *et al.* 1994; Karpychev & Fleitout 2000; Forte & Perry 2000; Panasyuk & Hager 2000; Forte & Mitrovica 2001; Pari 2001; Deschamps *et al.* 2001). This approach is useful when contending with significant uncertainties in the amplitudes of the seismic anomalies, as is evident in Fig. 11. This figure underscores a key concern, namely that the use of mineral physics data to translate the seismic anomalies can lead to large variations in the estimated mantle density anomalies, depending on the choice of tomography model. To avoid this difficulty, it is preferrable to carry out simultaneous inversions of all global geodynamic data (Fig. 2) to determine an optimal velocity-density scaling factor $\mathrm{dln}\,\rho/\mathrm{dln}\,V_S$ (e.g., Forte & Mitrovica 2001), such that:

$$\frac{\delta\rho}{\rho} = \left(\frac{\mathrm{dln}\,\rho}{\mathrm{dln}\,\mathrm{V_S}}\right)\frac{\delta V_S}{V_S}\,, \tag{103}$$

where $\delta V_S/V_S$ are shear velocity anomalies obtained from a tomography model and $\delta\rho/\rho$ are the relative perturbations in density which are constrained by the geodynamic data. This approach has also been employed in other tomography-based mantle flow models (e.g., Hager & Clayton 1989; Corrieu *et al.* 1994; Panasyuk & Hager 2000).

The geodynamic data provide direct, linear constraints on the density anomalies in the mantle (see equations 98 – 100), which may be generically expressed as follows:

$$\delta O_\ell^m = f_\ell \int_b^a K_\ell(\eta; r')\,(\rho_1)_\ell^m(r')\,dr' \approx f_\ell \sum_{i=1}^{N} K_\ell(\eta; r_i)\,(\rho_1)_\ell^m(r_i)\,w_i \tag{104}$$

in which δO_ℓ^m are the spherical harmonic coefficients of a geodynamic observable (e.g., geoid or free-air gravity anomalies), $K_\ell(\eta; r')$ is the corresponding kernel function and f_ℓ is a factor which depends on the geodynamic observable (see, for example, equation 98). The numerical calculation of integrals is usually carried out with equivalent finite sums, such as in (104) where w_i is a weighting term which depends on the numerical summation algorithm (e.g., Gauss-Legendre quadrature). On the basis of expression (103), we may further rewrite equation (104) as:

$$\delta O_\ell^m = f_\ell \sum_{i=1}^{N} w_i K_\ell(\eta; r_i)\,\rho_o(r_i)\left(\frac{\delta V_S}{V_S}\right)_\ell^m\!(r_i)\,\left(\frac{\mathrm{dln}\,\rho}{\mathrm{dln}\,\mathrm{V_S}}\right)\!(r_i) \tag{105}$$

in which $\delta V_S/V_S$ are the shear-velocity anomalies in the mantle and $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ is the corresponding velocity-to-density scaling which is assumed to vary with depth in the mantle.

Expression (105) provides the basis for a discrete, linear inversion of the geodynamic data δO_ℓ^m to find an optimal velocity-to-density scaling profile $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$. A series of Occam inversions (Constable *et al.* 1987) of the geodynamic data (Fig. 2) was carried out to find the smoothest family of $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ profiles which are consistent with the data. For each Occam inversion, a different seismic shear-wave tomography model (Fig. 11) was employed and in each case the same geodynamic kernel functions (Fig. 10) were employed to relate the mantle heterogeneity to the geodynamic data. The results of these inversions are summarised in Fig. 12 which shows the optimal profiles of $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$

and the resulting depth variation of the rms amplitude of the lateral density anomalies implied by expression (103).

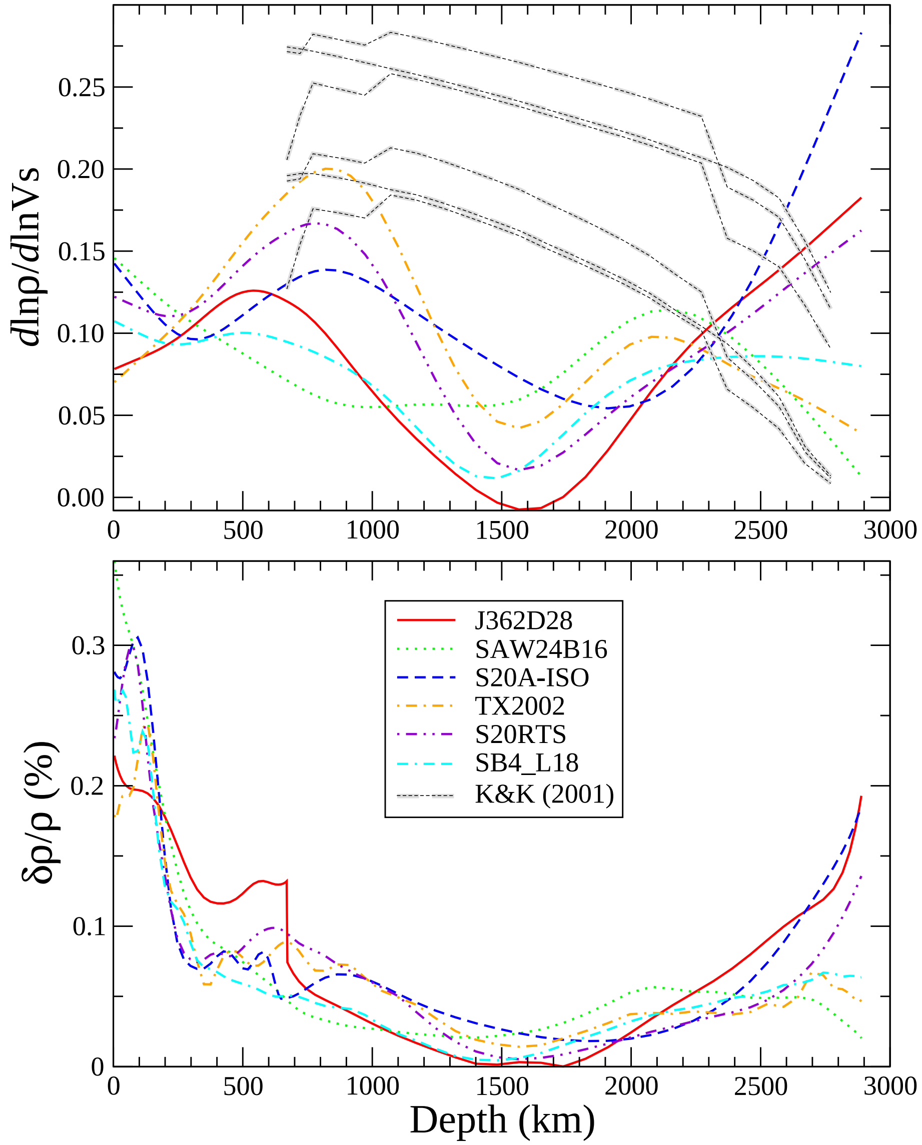

Figure 12: Occam inversions for mantle density anomalies. In the *top frame* are shown the optimal Occam inferred velocity-density scaling profiles for each of the tomography models illustrated in Fig. 11. The legend identifying the results is in the bottom frame. For comparison are shown (thick dashed gray lines) a series of theoretical estimates of the depth variation of $\mathrm{dln}\,\rho/\mathrm{dln}\,V_S$ in the lower mantle obtained by Karato & Karki (2001) on the basis of mineral physics data. These mineral physical estimates assume that both density and seismic anomalies are produced by temperature perturbations in the mantle. In the *bottom frame* are shown the resulting inferences of the depth variation of the rms density anomalies which are calculated on the basis of the seismic shear velocity anomalies using expression (103).

Mineral physics estimates of the velocity-density scaling (Fig. 12) show significant departures from the scaling profiles derived from the Occam inversions. As discussed above, this discrepancy may in part arise from the inadequate resolution of seismic heterogeneity in the mantle – especially in the mid-mantle region between 1000 km and 2000 km depth – and it may also arise from inadequacy of the theoretical assumptions inherent in the mineral physics estimates. In the latter case, the chief assumption which may be deficient is the dominance of thermal effects on both the seismic and density anomalies in the mantle. Compositional effects, especially on mantle density heterogeneity, may be one important explanation for the deviations from theoretical mineral physics estimates based on temperature effects alone (e.g., Stacey 1998). The relative importance of thermal and compositional contributions to mantle heterogeneity can be directly tested in the context of simultaneous inversions seismic and geodynamic data (Simmons *et al.* 2007, 2009, 2010) and these will be discussed below.

3.3 Predicted tectonic plate motions

In this and the next few sections we will consider the extent to which the tomography-based inferences of mantle density anomalies presented in Fig. 12 provide a successful explanation of the convection-related surface observables (Fig. 2). We begin here with a consideration of convection-induced tectonic plate motions. The predicted horizontal divergence of the plate motions, calculated on the basis of expresion (101) and using the kernels in Fig. 10b, are presented in Fig. 13 alongside the observed NUVEL-1 divergence of the plates. A detailed spectral comparison between the predicted and observed plate divergence is shown on Fig. 14.

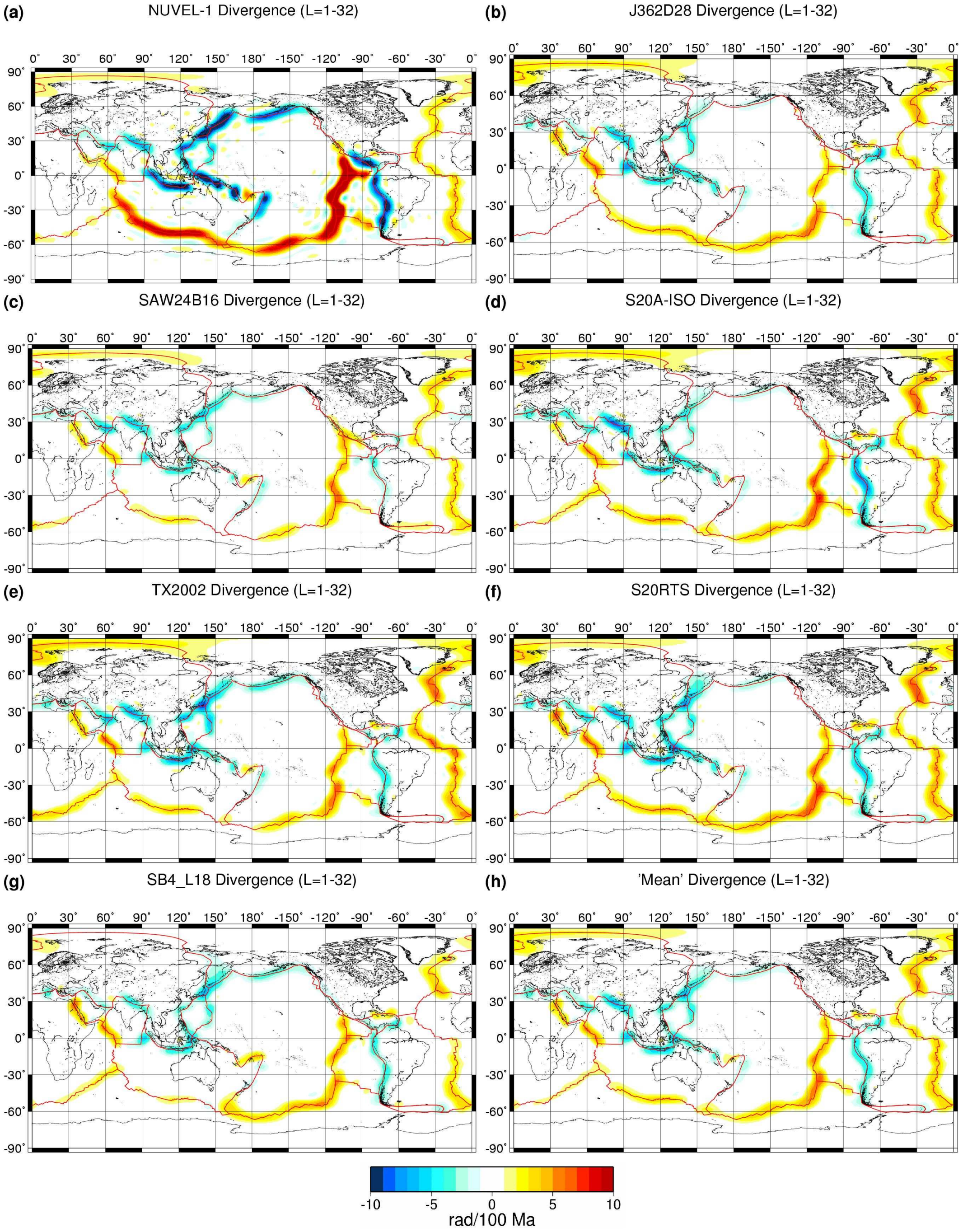
(a)
NUVEL-1 Divergence (L=1-32)
(b)
J362D28 Divergence (L=1-32)
(c)
SAW24B16 Divergence (L=1-32)
(d)
S20A-ISO Divergence (L=1-32)
(e)
TX2002 Divergence (L=1-32)
(f)
S20RTS Divergence (L=1-32)
(g)
SB4_L18 Divergence (L=1-32)
(h)
'Mean' Divergence (L=1-32)
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
-10 -5 0 5 10
rad/100 Ma

Figure 13: Tomography-based horizontal divergence of the tectonic plates. **(a)** Observed NUVEL-1 horizontal plate divergence (DeMets *et al.* 1990). **(b) – (g)** show the surface divergence predicted on the basis of tomography models *J362D28* (Antolik *et al.* 2003), *SAW24* (Mégnin & Romanowicz 2000), *S20A-Iso* (Ekström & Dziewonski 1998), *TX2002* (Grand 2002), *S20RTS* (Ritsema *et al.* 1999), *SB4_L18* (Masters *et al.* 2000), using the corresponding $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling profiles in Fig. 12. All shear-velocity heterogeneity models have been projected onto the common parametrisation in exptression (102) prior to calculating the predicted plate motions. **(h)** The 'mean' prediction of plate divergence obtained by calculating the statistical sample average of all the predictions (b) – (g). The observed and predicted plate motions shown here are all synthesized from spherical harmonics up to degree and order 32.

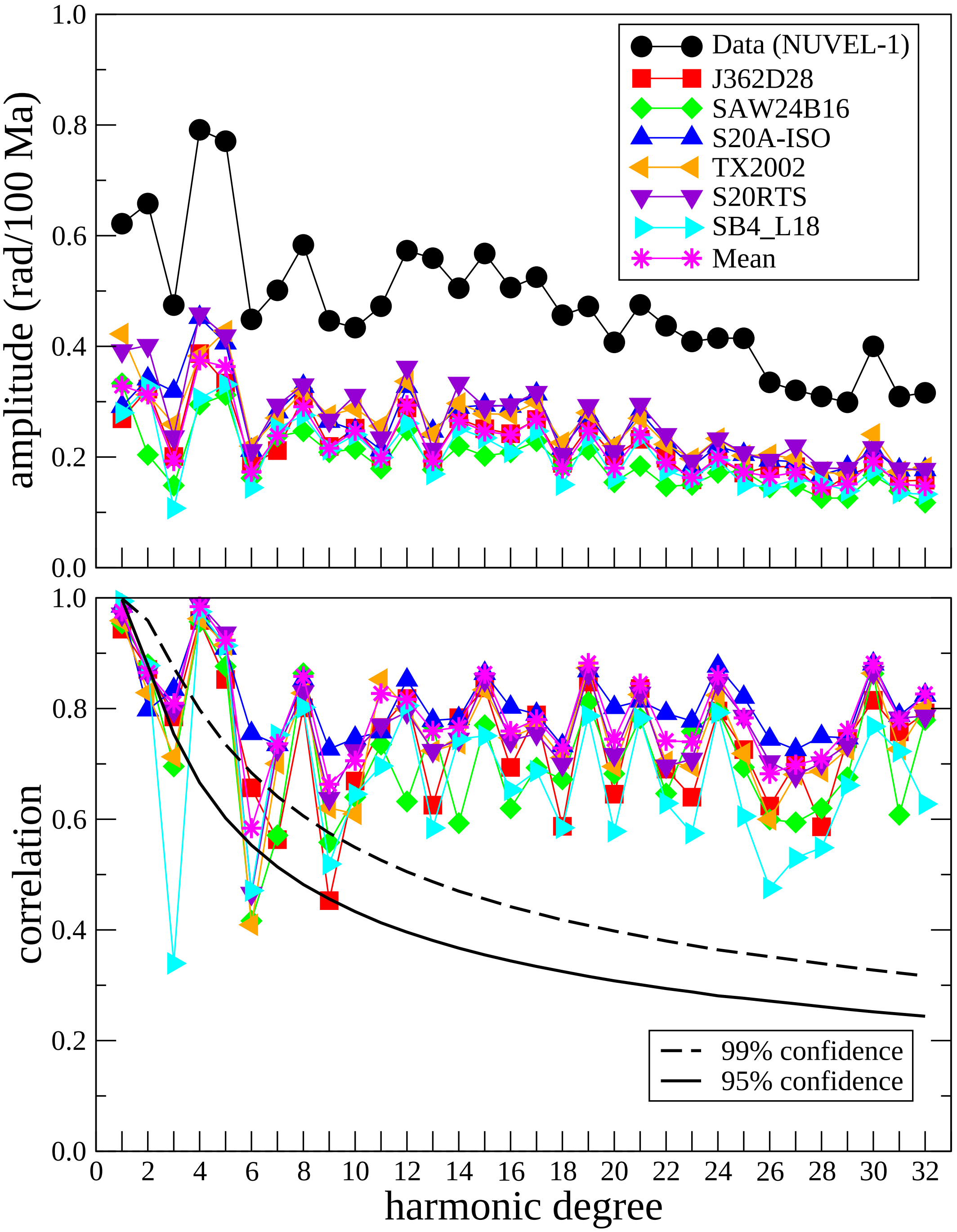

Figure 14: Spectral comparison of predicted and observed plate divergence. *Top frame* shows the degree variance, as defined in equation (3), of the observed and predicted plate divergence fields shown in Fig. 13. *Bottom frame* shows the degree correlations, as defined in equation (4). The black solid and dashed lines represent the 95% and 99% confidence levels, respectively.

A quantitative summary of the overall agreement between the predicted and observed plate divergence, quantified in terms of total rms amplitude, total spatial correlation and

Table 1: Comparison of observed† and predicted‡ horizontal plate divergence

Model	NUVEL*	*J362D28*	*SAW24*	*S20A-Iso*	*TX2002*	*S20RTS*	*SB4_L18*	*Mean*
rms	2.8	1.3	1.1	1.5	1.5	1.6	1.2	1.3
correl	–	0.77	0.74	0.83	0.78	0.80	0.74	0.82
var. red.	–	50%	44%	60%	56%	59%	45%	55%

Note: 'rms' = root-mean-square amplitude, expressed here in units of $10^{-8}a^{-1}$

'correl' = global cross-correlation, 'var. red.' = variance reduction

† Data and predictions (shown in Fig. 13) are synthesized from spherical harmonics up to degree 32.

‡ Predictions employ the tomography model indicated at the top of each column and they use the corresponding velocity-density scalings in Fig. 12.

* Horizontal divergence of the plates is derived from the NUVEL-1 (DeMets *et al.* 1990) description of present-day plate velocities.

variance reduction, is provided in Table 1. The variance reduction is defined as follows:

$$\text{var. red.} = 100\% \times \left[1 - \frac{\sum_\ell \sum_{m=-\ell}^{+\ell} (O-P)_\ell^{m*} (O-P)_\ell^m}{\sum_\ell \sum_{m=-\ell}^{+\ell} O_\ell^{m*} O_\ell^m}\right] \tag{106}$$

where $(O-P)_\ell^m$ and O_ℓ^m are the harmonic coefficients of the difference between the observed and predicted fields and the observed fields, respectively.

An important characteristic of the predicted divergence fields shown in Fig. 13, which is also confirmed in Table 1, is the subdued amplitude of the predicted fields relative to the data. As will be seen in the next few sections, all tomography-based predictions of geodynamic observables which are calculated on the basis of the optimal Occam-inverted density-velocity scaling profiles (Fig. 12) are systematically reduced in amplitude relative to the amplitude of the observed fields. This systematic bias to low predicted amplitudes is a natural consequence of the least-squares misfit criterion which is employed in almost all inversion algorithms (including the Occam approach) which leads to significantly damped inferences of the $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling profiles.

It can be readily shown that the inherent damping of tomography-based inferences of $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ arises as a consequence of poorly resolved 3-D seismic structure in geodynamically important regions of the mantle, such as the mid-mantle region (e.g., Forte *et al.* 1994). We can illustrate this in an inverse problem in which we seek, for the sake of simplicity, the best-fitting constant $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ value. As shown in expression (104), the predicted convection-related observables are linearly dependent on the internal density perturbations in the mantle. By virtue of expression (105), we may write

$$d(\theta,\varphi) = \left(\frac{\mathrm{dln}\,\rho}{\mathrm{dln}\,\mathrm{V_S}}\right)\, p(\theta,\varphi)\ ,$$

where $d(\theta,\varphi)$ represent the surface data and $p(\theta,\varphi)$ is the corresponding mantle-flow prediction assuming $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S} = 1$. We can then show that the optimal $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ value that minimises the least-squares misfit between data and the predictions is given by the following expression:

$$\frac{\mathrm{dln}\,\rho}{\mathrm{dln}\,\mathrm{V_S}} = \frac{\sigma_d}{\sigma_p}\, c(d,p)\ , \tag{107}$$

where σ_d and σ_p are the rms amplitudes of the data and the predictions,

$$\sigma_d^2 = \frac{1}{4\pi} \iint_{S_1} \mid d(\theta, \varphi) \mid^2 \; \mathrm{d^2S} \; , \; \sigma_\mathrm{p}^2 = \frac{1}{4\pi} \iint_{\mathrm{S_1}} \mid \mathrm{p}(\theta, \varphi) \mid^2 \; \mathrm{d^2S} \; ,$$

and

$$c(d, p) = \frac{1}{4\pi} \iint_{S_1} d(\theta, \varphi) p(\theta, \varphi) \; \mathrm{d^2S} \; , \tag{108}$$

where all integrals are defined over the surface of the unit sphere S_1. The quantity $c(d, p)$ defined in (108) is the cross-correlation between the data and the predictions. Equation (107) thus shows that the inferences of $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ will be directly proportional to the spatial correlation between the surface data and the corresponding geodynamic integral of the 3-D structure in the seismic tomography model. As previously shown by Forte *et al.* (1994) this correlation may be significantly degraded if the 3-D seismic structure is poorly resolved.

3.4 Predicted free-air gravity anomalies

Tomography-based mantle flow models of Earth's nonhydrostatic gravitational potential have almost exclusively been based on analyses of the equivalent geoid anomalies, beginning with the earliest studies by Richards & Hager (1984), Hager *et al.* (1985) and continuing to recent studies (e.g., Panasyuk & Hager 2000). Since the global geoid anomalies are strongly dominated by horizontal wavelengths corresponding to harmonic degrees $\ell = 2, \; 3$ they will effectively constrain only the longest wavelength components of 3-D mantle structure. To avoid this very-long wavelength bias we consider here a representation of the nonhydrostatic geopotential in terms of equivalent global free-air gravity anomalies (see expression 97). The free-air gravity anomalies contain a more evenly balanced representation of the different horizontal wavelengths in the nonhydrostatic geopotential (i.e., a 'flatter' amplitude spectrum).

The observed and predicted gravity anomaly fields, calculated on the basis of expressions (97, 98) and using the kernels in Fig. 10c,d, are presented in Fig. 15. A comparison of the predicted and observed gravity anomalies in the spectral domain is shown in Fig. 16. In Table 2 is presented a detailed summary of the global agreement between the predicted and observed gravity anomalies. In contrast to the predicted plate divergence discussed in the preceding section, we note that the amplitudes of the predicted gravity anomaly fields are somewhat less muted (compare Tables 1 and 2). The overall agreement with the gravity data (in terms of variance reduction) is however poorer than the plate divergence fits because of the decreased spatial correlation between the predictions and the data.

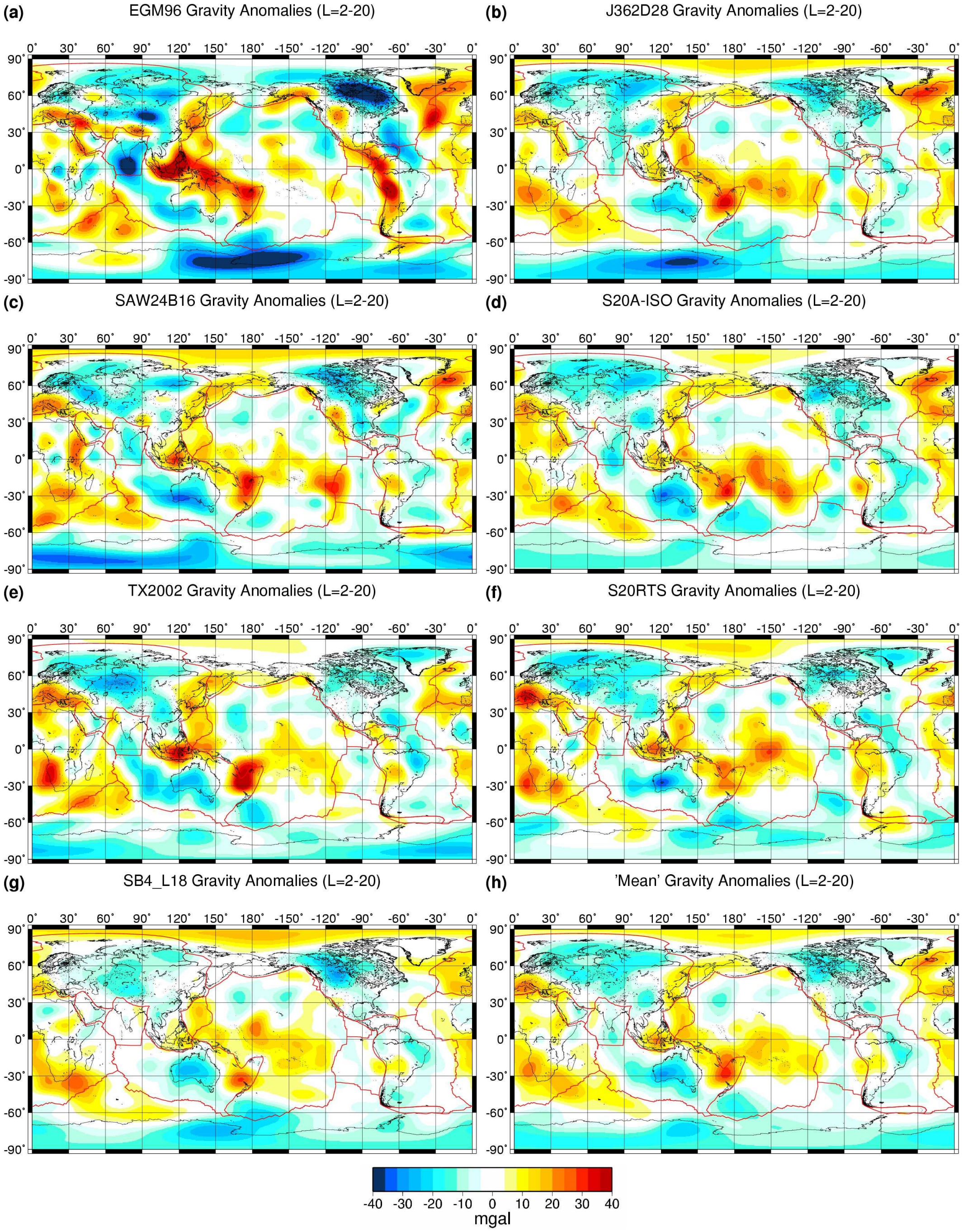
(a) EGM96 Gravity Anomalies (L=2-20)
(b) J362D28 Gravity Anomalies (L=2-20)
(c) SAW24B16 Gravity Anomalies (L=2-20)
(d) S20A-ISO Gravity Anomalies (L=2-20)
(e) TX2002 Gravity Anomalies (L=2-20)
(f) S20RTS Gravity Anomalies (L=2-20)
(g) SB4_L18 Gravity Anomalies (L=2-20)
(h) 'Mean' Gravity Anomalies (L=2-20)
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
-40 -30 -20 -10 0 10 20 30 40
mgal

Figure 15: Tomography-based global free-air gravity anomalies. **(a)** Observed EGM96 anomalous free-air gravity field (Lemoine *et al.* 1998). **(b) – (g)** show the free-air anomalies predicted on the basis of tomography models *J362D28* (Antolik *et al.* 2003), *SAW24* (Mégnin & Romanowicz 2000), *S20A-Iso* (Ekström & Dziewonski 1998), *TX2002* (Grand 2002), *S20RTS* (Ritsema *et al.* 1999), *SB4_L18* (Masters *et al.* 2000), using the corresponding $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling profiles in Fig. 12. All shear-velocity heterogeneity models have been projected onto the common parametrisation in exptression (102) prior to calculating the gravity anomalies. **(h)** The 'mean' free-air gravity anomaly field obtained by calculating the statistical sample average of all the predictions (b) – (g). The observed and predicted gravity anomalies shown here are all synthesized from spherical harmonics up to degree and order 20.

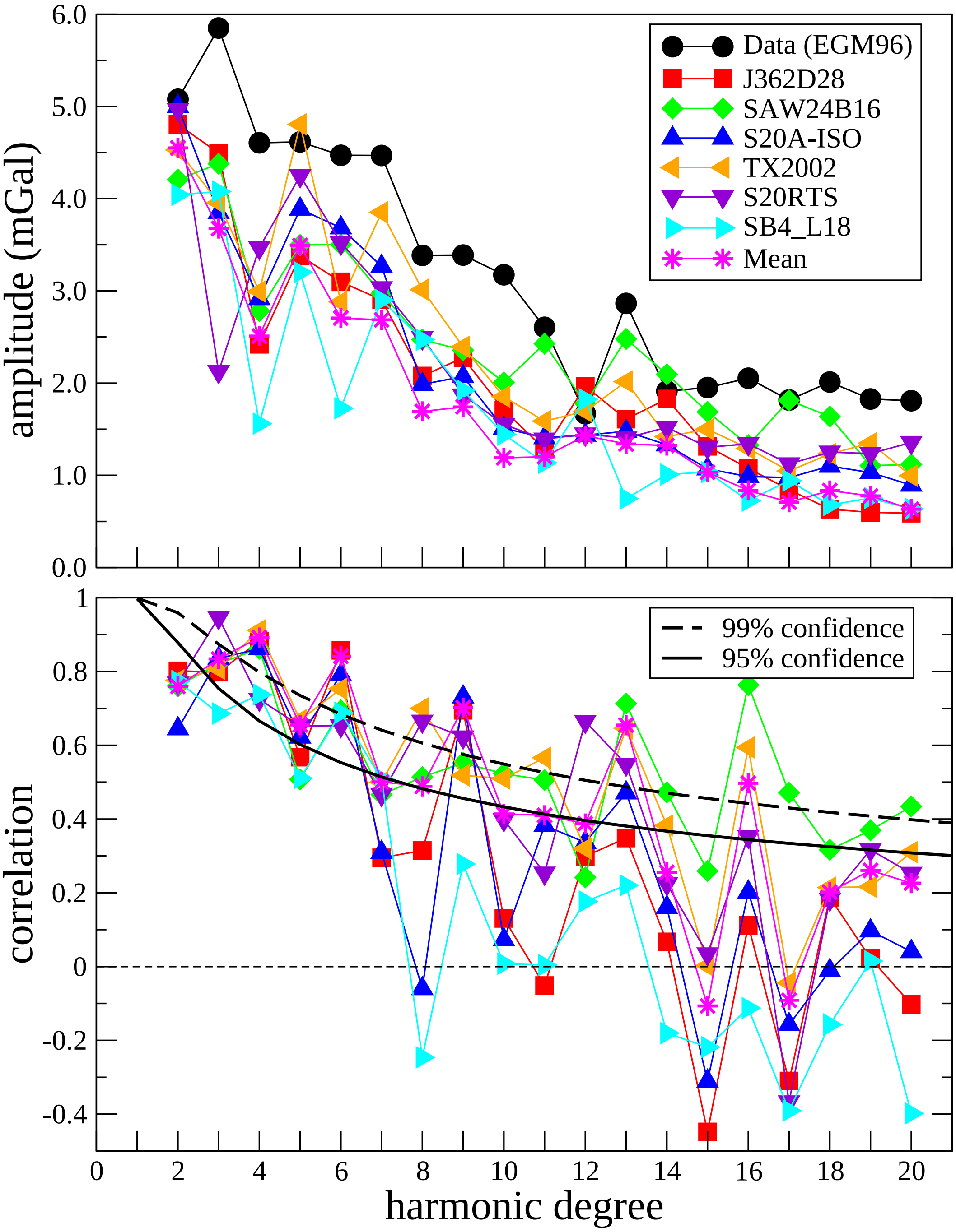

Figure 16: Spectral comparison of predicted and observed free-air gravity anomalies. *Top frame* shows the degree variance, as defined in equation (3), of the observed and predicted gravity anomaly fields shown in Fig. 15. *Bottom frame* shows the degree correlations, as defined in equation (4). The black solid and dashed lines represent the 95% and 99% confidence levels, respectively.

Table 2: Comparison of observed† and predicted‡ free-air gravity anomalies

Model	EGM96*	*J362D28*	*SAW24*	*S20A-Iso*	*TX2002*	*S20RTS*	*SB4_L18*	*Mean*
rms	14.8	10.4	11.2	10.6	11.4	10.5	8.9	9.3
correl	–	0.55	0.62	0.54	0.64	0.58	0.42	0.64
var. red.	–	28%	37%	26%	39%	33%	15%	41%

Note: 'rms' = root-mean-square amplitude, expressed here in units of $mGal$

'correl' = global cross-correlation, 'var. red.' = variance reduction

† Data and predictions (shown in Fig. 15) are synthesized from spherical harmonics up to degree 20.

‡ Predictions employ the tomography model indicated at the top of each column and they use the corresponding velocity-density scalings in Fig. 12.

* The global free-air gravity anomalies are derived from the EGM96 geopotential model (Lemoine *et al.* 1998).

3.5 Predicted dynamic surface topography

The term 'dynamic topography' is here defined to include all contributions to the topography of Earth's solid surface which arise from mantle convection. There has been some controversy as to origin and amplitude of Earth's dynamic topography, with some confusion as to how this dynamic topography should even be defined (e.g., Gurnis 1990; Forte *et al.* 1993b,c; Gurnis 1993). As will be discussed below, the interpretation of dynamic surface topography employed here, namely the topography arising from all density anomalies in the mantle (including those in the lithosphere), is not universally accepted. A detailed review and analysis of the opposing schools of thought in this debate has been presented by Pari (2001).

Numerous studies over the past three decades have focussed exclusively on the topography of ocean floor and the general conclusion is that this topography can be almost completely explained in terms of shallow, thermally induced density anomalies which arise from the age-dependent cooling of the oceanic lithosphere (e.g. Parson & Sclater 1977; Stein & Stein 1992). The oceanic bathymetry is modelled in terms of isostatic compensation of density anomalies in the oceanic lithosphere and hence many studies have regarded this form of thermal isostasy as not being a 'dynamic' contribution to surface topography. An adequate review of the extensive literature dealing with this interpretation of seafloor topography is well beyond the scope of this chapter, but recent studies of age-dependent oceanic bathymetry with extensive references to past analyses may be found in Doin & Fleitout (2000) and in Crosby et al. (2006).

Whether or not the thermal isostatic signal in seafloor bathymetry may be regarded as 'static' or 'dynamic' depends on the temporal variability of the structure of the upper thermal boundary layer (i.e., the lithosphere in oceanic regions) due to time-dependent mantle convection. A purely mechanical interpretation of the instantaneous, present-day isostatic compensation of thermal anomalies in the lithosphere ignores this background time-dependence. The lateral temperature variations in the cooling lithosphere are essentially maintained by a balance between vertical heat conduction and horizontal heat advection (e.g., Parsons & McKenzie 1978; Jarvis & Peltier 1982) and the latter is clearly a dynamic effect. In the absence of convection, the upper thermal boundary layer and the corresponding variations in oceanic bathymetry will vanish. The importance of time-dependent changes in the structure of the lithosphere, and hence the corresponding changes in the contribution to surface topography, is highlighted by the significant changes in plate tectonic velocities and plate geometries in the Cenozoic (e.g., Lithgow-Bertelloni & Richards 1998) which suggest significant departures from steady-state conditions.

In the absence of convection, the only contribution to surface topography arises from the isostatic compensation of lateral variations in crustal thickness and crustal density. Earth's present-day topography is therefore the superposition of the crustal isostatic topography and the dynamic topography – as defined here – due to density anomalies in the mantle which are maintained by mantle convection (Forte *et al.* 1993b). The crustal isostatic topography is large and it explains most of the observed present-day topography on Earth (e.g., Forte & Perry 2000; Pari & Peltier 2000). The dynamic topography is obtained by subtracting the crustal isostatic topography from the observed topography and this therefore requires an accurate model of crustal structure. This requirement presents a major challenge because of the uneven and incomplete seismic sampling global crustal thickness and also because of the significant uncertainties in constraining the density of the crust. Currently, the most complete compilation of global crustal heterogeneity is model CRUST2.0 (Bassin *et al.* 2000). This crustal model was employed in retrieving the dynamic surface topography presented above in Fig. 2b.

The observed and predicted dynamic topography fields, calculated on the basis of expresions (99) and using the kernels in Fig. 10e,f, are presented in Fig. 17. A quantitative spectral comparison of the predicted and observed dynamic topography is shown in Fig. 18. In Table 3 is presented a detailed summary of the match between the predicted and observed topography fields. The topography kernels (Fig. 10e,f) show that near-surface density anomalies provide the strongest contributions to the predicted surface topography (Fig. 17) and that they will be in near-isostatic equilibrium, where perfect isostatic compensation corresponds to a kernel value of -1. We note that in the case of density anomalies which effectively see a no-slip surface boundary (given by expression 92), the condition of near-isostasy extends to depths of about 200 km (see Fig. 10f).

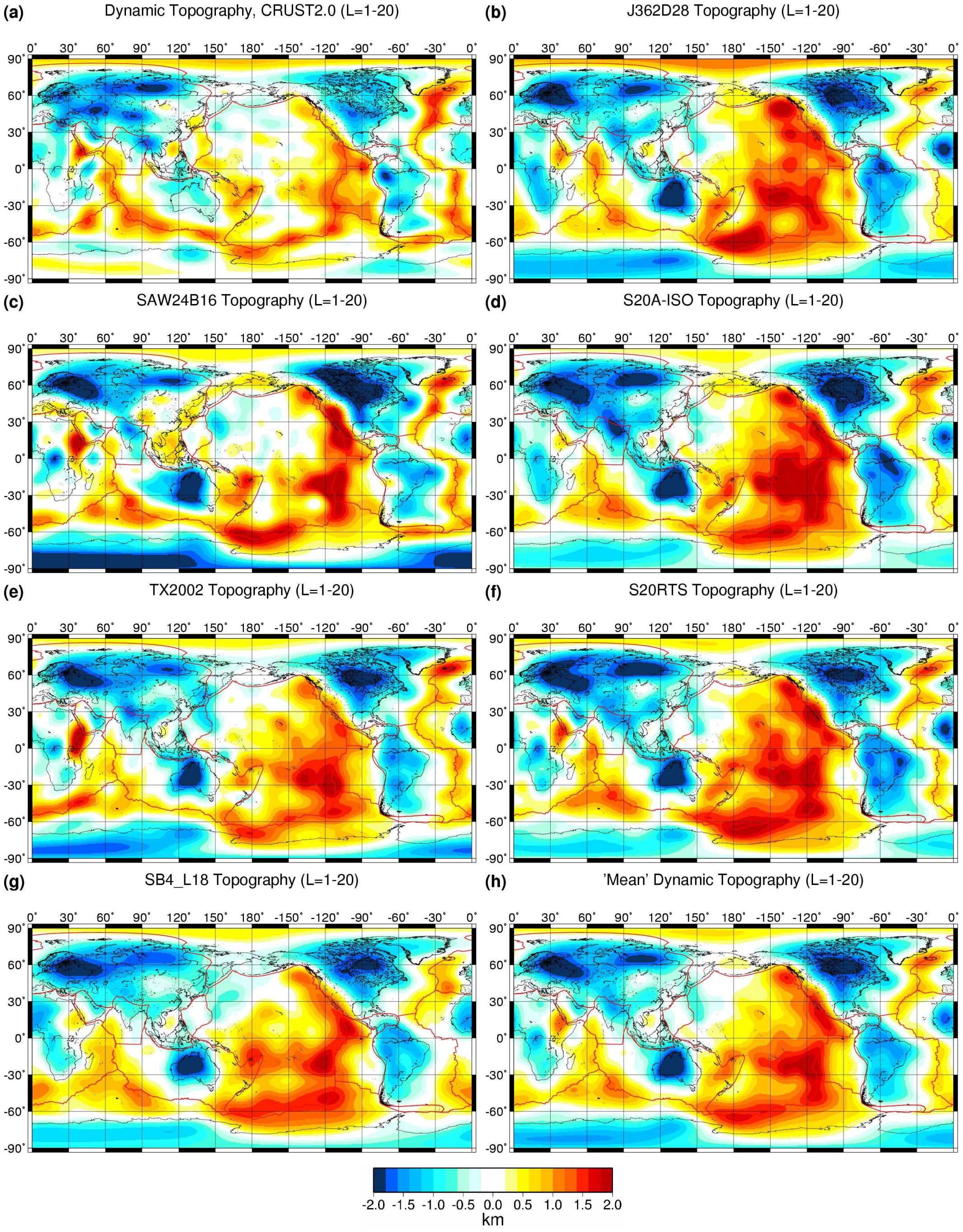
(a) Dynamic Topography, CRUST2.0 (L=1-20)
(b) J362D28 Topography (L=1-20)
(c) SAW24B16 Topography (L=1-20)
(d) S20A-ISO Topography (L=1-20)
(e) TX2002 Topography (L=1-20)
(f) S20RTS Topography (L=1-20)
(g) SB4_L18 Topography (L=1-20)
(h) 'Mean' Dynamic Topography (L=1-20)
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
-2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0
km

Figure 17: Tomography-based global dynamic surface topography. **(a)** Dynamic surface topography estimated by removing isostatically compensated crustal heterogeneity – described by model CRUST2.0 (Bassin *et al.* 2000) – from the observed present-day surface topography. **(b) – (g)** show the dynamic surface topography predicted on the basis of tomography models *J362D28* (Antolik *et al.* 2003), *SAW24* (Mégnin & Romanowicz 2000), *S20A-Iso* (Ekström & Dziewonski 1998), *TX2002* (Grand 2002), *S20RTS* (Ritsema *et al.* 1999), *SB4_L18* (Masters *et al.* 2000), using the corresponding $\mathrm{d}\ln\rho/\mathrm{d}\ln \mathrm{V_S}$ scaling profiles in Fig. 12. All shear-velocity heterogeneity models have been projected onto the common parametrisation in exptression (102) prior to calculating the dynamic topography. **(h)** The 'mean' dynamic surface topography field obtained by calculating the statistical sample average of all the predictions (b) – (g). The crust-corrected and predicted dynamic topography shown here are all synthesized from spherical harmonics up to degree and order 20.

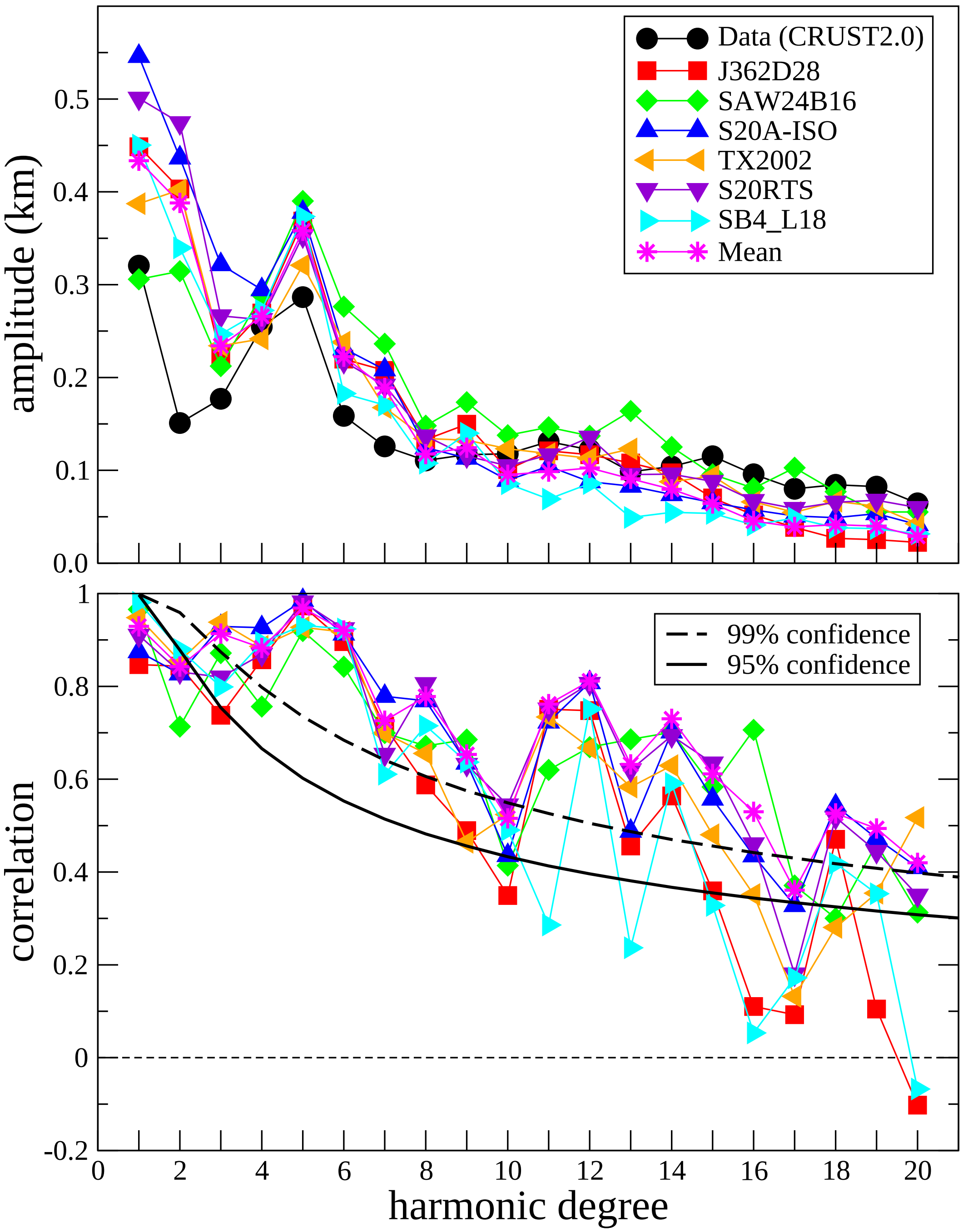

Figure 18: Spectral comparison of predicted and observed (crust-corrected) dynamic surface topography. *Top frame* shows the degree variance, as defined in equation (3), of the observed and predicted topography fields shown in Fig. 17. *Bottom frame* shows the degree correlations, as defined in equation (4). The black solid and dashed lines represent the 95% and 99% confidence levels, respectively.

In view of the debate concerning the magnitude of the thermal isostatic contributions to surface topography in oceanic regions, it is helpful to explore the relative importance of shallow (i.e. lithospheric) and deep-mantle (i.e. sub-lithospheric) buoyancy using the

Table 3: Comparison of observed† and predicted‡ dynamic surface topography

Model	CRUST*	*J362D28*	*SAW24*	*S20A-Iso*	*TX2002*	*S20RTS*	*SB4_L18*	*Mean*
rms	696	908	893	1002	860	972	849	873
correl	–	0.75	0.76	0.80	0.78	0.78	0.79	0.87
var. red.	–	25%	30%	22%	41%	24%	44%	47%

Note: 'rms' = root-mean-square amplitude, expressed here in units of m

'correl' = global cross-correlation, 'var. red.' = variance reduction

† Data and predictions (shown in Fig. 17) are synthesized from spherical harmonics up to degree 20.

‡ Predictions employ the tomography model indicated at the top of each column and they use the corresponding velocity-density scalings in Fig. 12.

* The global dynamic surface topography obtained by subtracting CRUST2.0 (Bassin *et al.* 2000) isostatic crust from the observed topography.

seismic tomography models. In this regard it is useful to first evaluate the extent to which the tomography-based topography predictions incorporate the age-dependent cooling of the oceanic lithosphere. Tomography models employing seismic travel-time and/or surface wave data place strong constraints on lithospheric mantle heterogeneity and they contain a clear signature of the oceanic cooling history (e.g., Woodward & Masters 1991; Su et al. 1992; Trampert & Woodhouse 1996; Ritzwoller et al. 2004).

Here we consider the global tomography models *SAW24* (Mégnin & Romanowicz 2000) and *TX2002* (Grand 2002), employed in Figs. 17c and 17e respectively. The age-dependent lateral temperature variations in the oceanic lithosphere were calculated using a simple halfspace cooling model with the same thermal parameters as the $GDH1$ plate-cooling model of Stein & Stein (1992). The seafloor age was determined using the Muller et al. (1997) digital isochrons. The global ocean basins were then sampled on a $1^\circ \times 1^\circ$ grid and for each grid cell the corresponding values of ocean age, shear velocity perturbation dV_S/V_S and temperature perturbation ΔT were determined. These samples were then averaged into two million-year age bins and the resulting relationship between seismic shear velocity and temperature is plotted in Fig. 19.

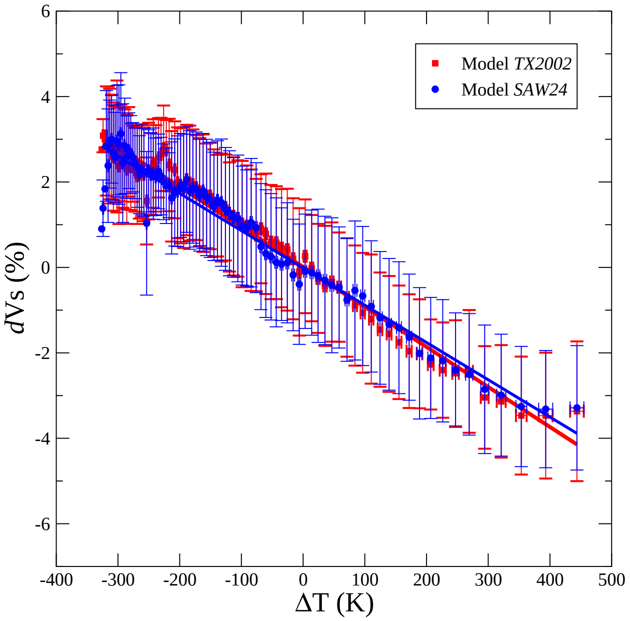

Figure 19: Comparison of seismic shear velocity and temperature anomalies in the oceanic lithosphere. Seismic shear velocity anomalies, vertically averaged in the depth interval 0-100km, are derived from tomography models *SAW24* (Mégnin & Romanowicz 2000) and *TX2002* (Grand 2002). The S-velocity anomalies, measured relative to the global oceanic mean value, were averaged into 2-Ma age bins using the global ocean-age compilation of Muller et al. (1997). The age-dependent temperature variations in the oceanic lithosphere were determined using a simple halfspace cooling model with the same thermal parameters as the $GDH1$ plate model of Stein & Stein (1992). These temperature variations were vertically averaged down to 100 km depth and subsequently averaged into 2-Ma age bins. The temperature anomalies are measured relative to the global oceanic mean value. The vertical and horizontal error bars represent 1 standard deviation relative to the average shear-velocity and temperature, respectively, in each age bin. The solid blue and red lines are the best-fitting linear regression lines for models *SAW24* and *TX2002*, respectively.

A linear regression analysis (solid lines in Fig. 19) yields an excellent fit to the binned $V_S - T$ variation and the slopes of the regression lines provide estimates of the effective temperature derivative of S-velocity in the depth range 0-100 km. These estimated derivatives are 8.7×10^{-5} K^{-1} and 9.4×10^{-5} K^{-1} for models *SAW24* and *TX2002*, respectively. These effective thermal derivatives agree well with independent mineral physics values determined by Stixrude & Lithgow-Bertelloni (2005) which range between 8 and 9 $\times 10^{-5}$ K^{-1}. These results confirm that the global tomography models considered in this study successfully resolve the pattern and amplitude of age-dependent cooling of the oceanic lithosphere.

The relative importance of surface topography due to lithospheric versus sub-lithospheric density anomalies is illustrated in Fig. 20. As expected, the topography contributions in oceanic regions from density anomalies in the upper 200 km of the mantle are dominated by the age-dependent cooling of the oceanic lithosphere. Depressions in continental regions exceed observational estimates (Fig. 17a) because the use of a purely depth dependent $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling factor (Fig. 12) does not resolve the intrinsic chemical buoyancy in the sub-continental mantle. This intrinsic buoyancy can be included in geodynamic models which allow for lateral variations of $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ in the shallow mantle (e.g., Forte & Perry 2000; Deschamps et al. 2001).

The topography generated by sub-lithospheric buoyancy (Fig. 20b), with rms amplitude equal to 410 metres, is comparable to that generated by lithospheric density anomalies (Fig. 20a) which has an rms amplitude equal 580 metres. These sub-lithospheric contributions to topography in the central Pacific and Atlantic oceans, and in southern Africa, are surface expressions of deep-seated buoyancy with sources in the lower mantle (Fig. 3). The depressions below the eastern and western margins of the Pacific (Fig. 20b) may be interpreted in terms of present-day and Cenozoic subduction history (e.g. Lithgow-Bertelloni & Gurnis 1997).

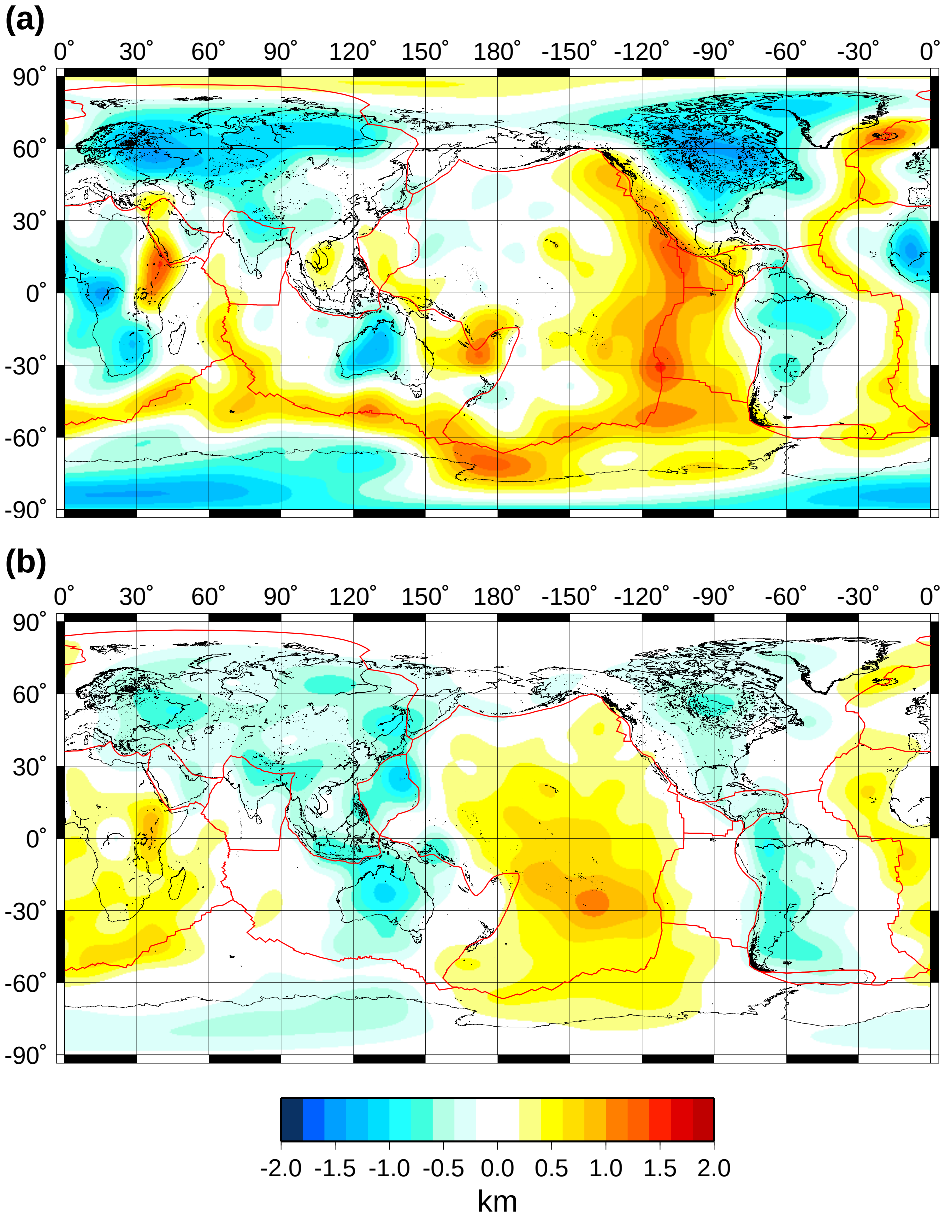
(a)
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
(b)
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
-2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0
km

Figure 20: Shallow- versus deep-mantle contributions to surface topography. **(a)** Surface topography predicted on the basis of tomography-derived density anomalies in the upper 200 km of the mantle. **(b)** Surface topography predicted on the basis of all density anomalies below 200 km depth. In both cases the density anomalies are derived from tomography model *TX2002* (Grand 2002) and using the corresponding $\mathrm{d}\ln\rho/\mathrm{d}\ln \mathrm{V_S}$ scaling profiles in Fig. 12. The summation of the topography fields in maps (a) and (b) yields the total field shown above in Fig. 17e. All fields are synthesized from spherical harmonics up to degree and order 20.

We note from Fig. 17 that the amplitude of the predicted dynamic topography is somewhat larger than the crust-corrected estimate of the observed dynamic topography. Table

3 shows that the rms amplitude of the mean topography prediction is 25% larger than the observed topography. This discrepancy is in marked contrast to that of the predicted gravity and plate divergence fields (Tables 1,2) which are, on average, 40-50 % smaller than the corresponding observed fields. The mismatched amplitudes of the predicted surface topography has motivated previous efforts (e.g., Le Stunff & Ricard 1995, 1997) to introduce additional buoyancy forces in the mantle arising, for example, from undulations of the 670 km seismic discontinuity (e.g., Thoraval et al. 1995; Forte & Woodward 1997b) which internally compensate the excess surface topography.

In assessing the significance of this discrepancy, in terms of mantle dynamics, it is important to recognise that the crustal correction employed in Fig. 17a contains significant uncertainties which may in part be responsible for the mismatched amplitudes. Perhaps more important is the question of the validity of a purely depth-dependent velocity-density scaling (Fig. 12) employed in the topography predictions. Such depth-dependence, which may be justified if one assumes that thermal effects on mantle heterogeneity dominate (e.g., Karato & Karki 2001), is not a good approximation in the presence of large lateral variations in chemical composition. Such compositional heterogeneity is expected to be important in the shallow subcontinental mantle and it will oppose the local thermal buoyancy, implying significant reductions in the amplitude of continental surface topography (e.g. Forte & Perry 2000).

To understand why the gravity (and plate motion) predictions, in contrast to the surface topography, have substantially smaller amplitudes than the observed fields, we must first appreciate that the global seismic data have difficulty resolving heterogeneity in the deep mantle, particularly in the southern hemisphere (e.g., Forte et al. 1994; Simmons et al. 2006). This difficulty does not apply to the shallow mantle where, as shown above (Fig. 19), the tomography models are able to successfully resolve the cooling oceanic lithosphere. The predicted surface topography is dominated by the density anomalies located in the upper mantle, whereas the gravity predictions contain significant contributions from lower-mantle density anomalies (Fig. 10). Joint inversions of global seismic and convection-related data sets have shown that increased amplitudes of the predicted gravity anomalies and substantially improved fits to the global gravity data can be achieved by modifying deep-mantle heterogeneity in a way that is consistent with both the seismic and geodynamic constraints (e.g., Forte et al. 1994; Simmons et al. 2006, 2007, 2009, 2010). Such modifications to lower-mantle heterogeneity have little impact on the amplitude of the predicted surface topography. These basic observations suggest that current mismatches between the amplitudes of the predicted and observed dynamic topography are not more significant, or fundamentally different, from the mismatch between the other predictions (i.e., gravity and plate motions) and their corresponding observables.

3.6 Predicted CMB topography

Tomography-based studies of the flow-induced deformations of the core-mantle boundary (CMB) has been rather limited (e.g., Hager *et al.* 1985; Forte *et al.* 1993a, 1995) relative to the much greater number of studies which have focussed on the surface geoid or topography. This is in large part a consequence of the uncertain and often contrasting inferences of the CMB topography which have been derived from inversions of seismic phases which are sensitive to this topography (e.g., Morelli & Dziewonski 1987; Rodgers & Wahr 1993, Obayashi & Fukao 1997; Boschi & Dziewonski 2000; Garcia & Souriau 2000; Sze & van der Hilst 2003). A detailed discussion of the seismological uncertainties and difficulties in obtaining reliable global images of the CMB topography may be found in Garcia & Souriau (2000).

The most accurate constraint on the topographic undulations of the CMB are derived from space-geodetic analyses of Earth's nutations. A brief review of these constraints was presented above in section 2.1. Unfortunately, these geodetic analyses can only constrain the excess flattening or dynamic ellipticity of the CMB – albeit with high precision – and this corresponds to only one coefficient ($\ell = 2,\ m = 0$) in a full spherical harmonic expansion (expression 1) of the spatially varying CMB topography field. This constraint on excess CMB flattening is included in all the Occam-inferred $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling profiles in Fig. 12.

For the purpose of comparison we consider a relatively recent seismological inference of the CMB topography (Fig. 21a) obtained by Boschi & Dziewonski (2000) (henceforth referred to as 'BD2000') on the basis of seismic PcP and PKP travel-time delays. The flow-induced CMB topography predicted on the basis of the Occam-inferred $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ profiles (Fig. 12), and using the kernels in Fig. 10g,h, is presented in Fig. 21. All predictions of CMB topography shown here exactly reproduce the geodetically constrained $400m$ excess CMB ellipticity. A quantitative spectral analysis (Fig. 22) of the predicted CMB topography fields shows that they are strongly dominated by the horizontal wavelengths corresponding to harmonic degrees $\ell = 2,\ 3$. In Table 4 is presented a detailed summary of the match between the predicted and observed CMB topography fields.

The overall spatial correlation between the predicted maps of flow induced CMB topography and the BD2002 model (Fig. 21) is quite good, however it is quite clear that there is a significant difference in the overall amplitude. The CMB topography obtained on the basis of the tomography-based flow calculations is approximately a factor of 3 to 4 times smaller than that in BD2002. Although, as noted above (section 3.3), there likely exists an inherent damping in the flow predictions, it is not possible to increase the amplitude of the predicted CMB topography (by increasing, for example, the magnitude of the $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling) without at the same time significantly degrading the fit to the other geodynamic observables. The $\pm 1.5\ km$ scale undulations in the predicted, flow-induced CMB topography are in the range estimated by Garcia & Souriau (2000) for horizontal wavelengths in excess of 1200 km.

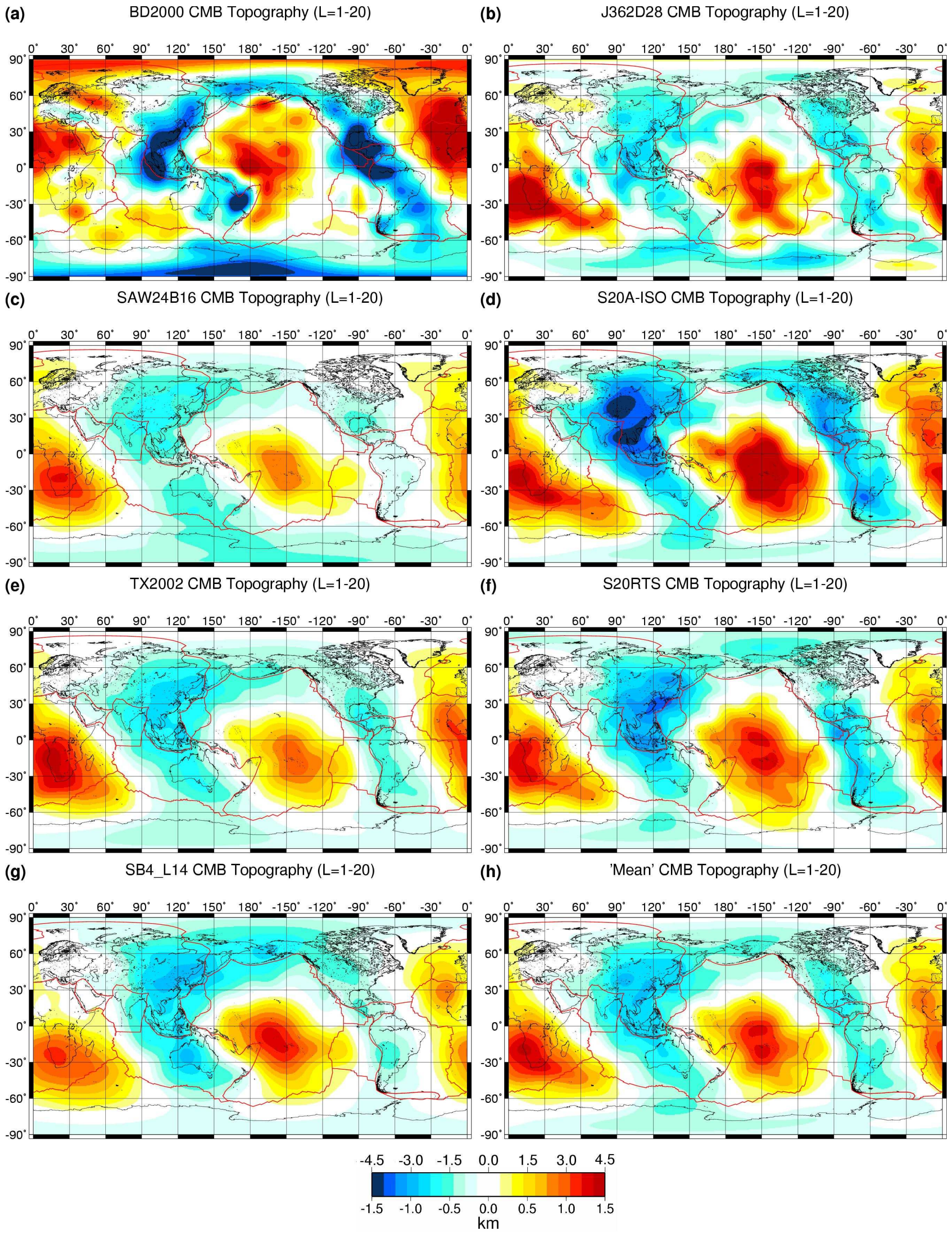
(a)
BD2000 CMB Topography (L=1-20)
(b)
J362D28 CMB Topography (L=1-20)
(c)
SAW24B16 CMB Topography (L=1-20)
(d)
S20A-ISO CMB Topography (L=1-20)
(e)
TX2002 CMB Topography (L=1-20)
(f)
S20RTS CMB Topography (L=1-20)
(g)
SB4_L14 CMB Topography (L=1-20)
(h)
'Mean' CMB Topography (L=1-20)
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
-4.5 -3.0 -1.5 0.0 1.5 3.0 4.5
-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5
km

Figure 21: Tomography-based global CMB topography. **(a)** Seismically inferred CMB topography derived by Boschi & Dziewonski (2000). The amplitudes correspond to the top values ($\pm 4.5\ km$) in the bottom scale bar. **(b) – (g)** show the CMB topography predicted on the basis of tomography models *J362D28* (Antolik *et al.* 2003), *SAW24* (Mégnin & Romanowicz 2000), *S20A-Iso* (Ekström & Dziewonski 1998), *TX2002* (Grand 2002), *S20RTS* (Ritsema *et al.* 1999), *SB4_L18* (Masters *et al.* 2000), using the corresponding $\mathrm{dln}\,\rho/\mathrm{dln}\,\mathrm{V_S}$ scaling profiles in Fig. 12. All shear-velocity heterogeneity models have been projected onto the common parametrisation in exptression (102) prior to calculating the CMB topography. **(h)** The 'mean' CMB topography field obtained by calculating the statistical sample average of all the predictions (b) – (g). The amplitudes for (b) – (h) correspond to the bottom values ($\pm 1.5\ km$) in the bottom scale bar. The observed and predicted CMB topography shown here are all synthesized from spherical harmonics up to degree and order 20.

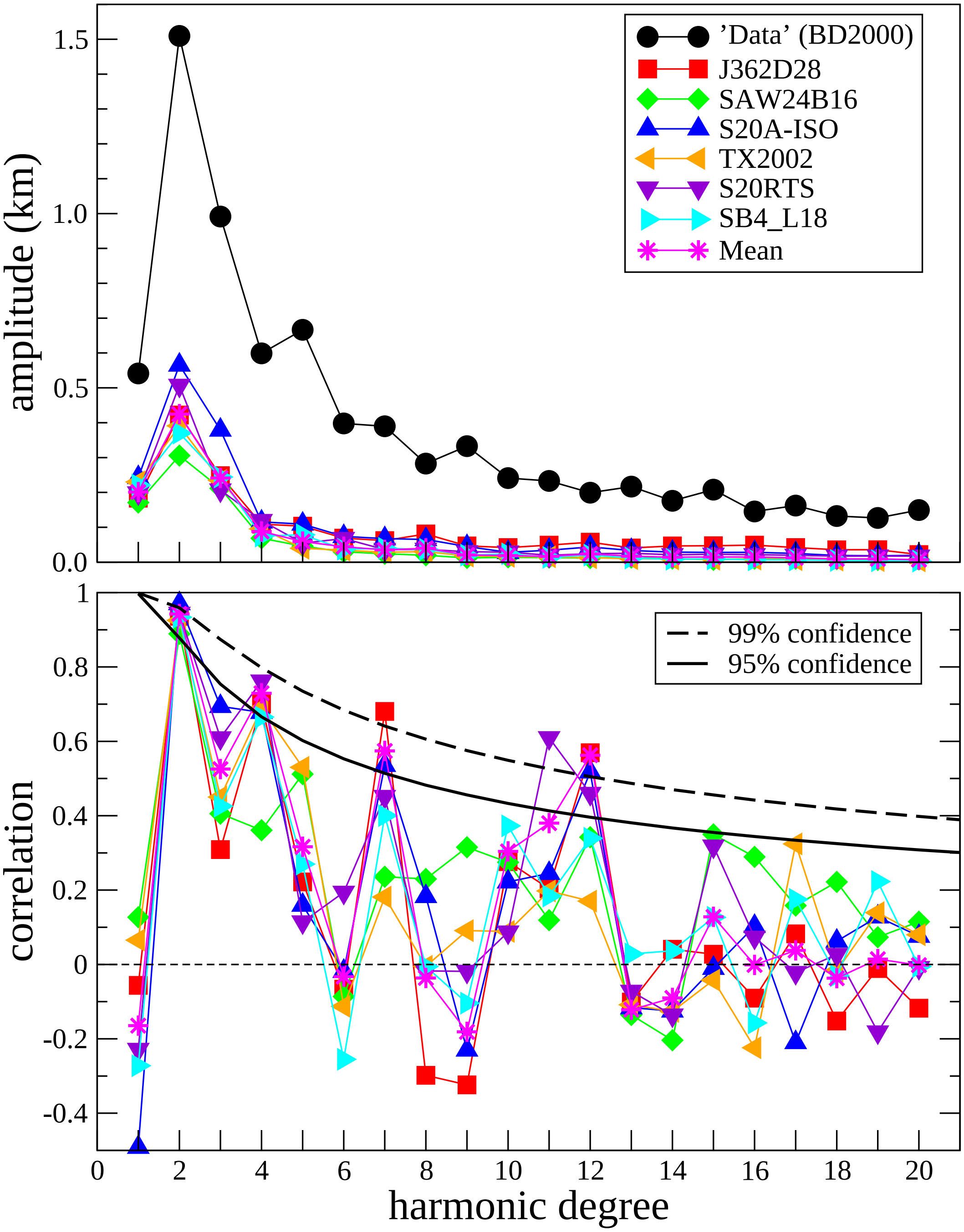

Figure 22: Spectral comparison of predicted and observed CMB topography. *Top frame* shows the degree variance, as defined in equation (3), of the observed and predicted CMB topography fields shown in Fig. 21. *Bottom frame* shows the degree correlations, as defined in equation (4). The black solid and dashed lines represent the 95% and 99% confidence levels, respectively.

Table 4: Comparison of seismically observed† and predicted‡ CMB topography

Model	BD2000*	*J362D28*	*SAW24*	*S20A-Iso*	*TX2002*	*S20RTS*	*SB4_L18*	*Mean*
rms	2.3	0.58	0.42	0.76	0.53	0.61	0.51	0.55
correl	–	0.56	0.56	0.64	0.59	0.65	0.55	0.62
var. red.	–	22%	18%	31%	22%	27%	20%	24%

Note: 'rms' = root-mean-square amplitude, expressed here in units of km

'correl' = global cross-correlation, 'var. red.' = variance reduction

† Seismic observations and predictions (shown in Fig. 21) are synthesized from spherical harmonics up to degree 20.

‡ Predictions employ the tomography model indicated at the top of each column and they use the corresponding velocity-density scalings in Fig. 12.

* CMB topography derived by Boschi & Dziewonski (2000) from global tomographic inversion of seismic delay times.

4 Joint seismic-geodynamic inversions for 3-D structure and flow in the mantle

In the preceding sections all interpretations of the 3-D structure and flow in the mantle, and the predicted surface geodynamic observables associated with this flow, have been entirely based on purely seismic inversions for 3-D lateral heterogeneity in the mantle. This approach, in which previously published seismic tomography models are converted *a posteriori* into an equivalent mapping of 3-D density anomalies, as described above in section 3.2, is followed in essentially all analyses of tomography-based mantle convection that have been published since the advent of the first global tomography models in the mid 1980's (e.g., Forte & Peltier 1987) up to the present (e.g., Conrad *et al.* 2011, Becker & Faccenna 2011, Alisic *et al.* 2012).

In the following we will instead consider how the seismic constraints on global heterogeneity in the mantle may be simultaneously inverted with independent constraints on 3-D structure from geodynamic and mineral physical data to obtain a new generation of mantle tomography models that successfully reconcile all these constraints, thereby providing new insights in the structure and dynamics of the mantle convective flow.

4.1 Deficiencies in mantle flow models based on pure seismic tomography inversions

While the *a posteriori* scaling of mantle tomography models, derived purely from seismic data, into equivalent density heterogeneity models has provided reasonably good fits to the global convection-related data sets, there remain significant unexplained misfits to these surface geodynamic constraints. More specifically, the variance reductions achieved in fitting long-wavelength global plate motions, gravity and dynamic topography anomalies are 55%, 41% and 47%, respectively. These misfits characterize the summary or '*Mean*' predictions in the final column of Tables 1,2,3. These results, obtained with optimized depth-dependent velocity-density scaling factors, suggest that, on average, about 50% of the power in convection-related data cannot be explained by the *a posteriori* scaling of the currently available seismic tomography models.

In assessing the origin of the outstanding misfit between the tomography-based predictions of the convection-related observables and the corresponding data, a number of possible deficiencies in the mantle-flow calculations should be considered. As discussed above, in section 2.3.1, a major simplification in the mantle flow modelling is the assumption that the sub-lithospheric viscosity of the mantle varies with depth only. We should note, however, that these flow models incorporate the effect of what is arguably the greatest-amplitude lateral heterogeneity of rheology in the solid-Earth: the tectonic plates. All flow predictions presented in the preceding sections are carried out with surface tectonic plates coupled to the mantle flow, and driven by this flow. The outstanding question is therefore whether lateral viscosity variations (LVV) in the underlying mantle can account for a significant fraction of the current misfit to the geodynamic data.

In addressing the impact of LVV on global convection-related surface data, *Moucha et al.* [2007] found that global-scale heterogeneity in viscosity, in which the amplitude of the LVV is equivalent to the variations in the mean depth-dependent viscosity (up to 3 orders of magnitude), yields predictions that differ by about 20% from those obtained by neglecting LVV. These differences were also found to be substantially smaller than the differences obtained using two different global tomography models. These results suggest that LVV in the mantle are not likely to be the principal source of the outstanding 50% misfit between the current tomography-based flow predictions and the surface geodynamic observables.

Another, possibly major, source of misfit may be due to the neglect of lateral variations in the scaling between seismic velocity and density, as would be the case in the presence of significant compositional heterogeneity in the mantle. The predicted geodynamic observables presented above in section 3 are obtained using a purely depth-dependent scaling, as expected if thermal contributions to mantle heterogeneity are assumed dominant (see, for example, Fig. 12).

A detailed exploration of the amplitude and distribution of chemical heterogeneity inferred on the basis of previously published tomography models, to achieve maximum agreement with the surface geodynamic data, was carried out by Forte (2007). This analysis allowed for explicit lateral variations in the velocity-to-density scaling by inverting for different scaling factors in regions of the mantle characterised by anomalously hot material (e.g. rising plumes) and anomalously cold material (e.g. subducting lithosphere in the lower mantle, and deep subcontinental roots in the upper mantle). This paramaterisation of chemical heterogeneity is analogous to that advocated by Xu *et al.* (2008), in which the composition of the mantle is assumed to be a mechanical mixture of basalt and harzburgite generated by a long-term history of partial melting at mid-ocean ridges and subsequent transport into the deep mantle by subducting lithosphere. The inversions by Forte (2007) for a laterally variable scaling, associated with 'hot' and 'cold' regions of the mantle, yielded an improved fit to the surface convection-related data sets, but the improvement of about 10% is largely insufficient to account for the outstanding 50% misfit to these data.

We refer the interested reader to Forte (2007) for a detailed discussion of the geodynamic and mineral physical interpretation of global tomography models, derived solely from seismic data, in terms of compositional heterogeneity in the mantle. An alternative view of compositional heterogneity in the mantle, based on 'probabilistic' (purely seismic) inversions for 3-D mantle structure is presented in Mosca *et al.* (2012).

4.2 Mantle flow modelling based on tomography models derived from seismic data alone is insufficient

The considerations presented in the previous section suggest that the origin of the outstanding (nearly 50%) misfit between the tomography-based predictions of the convection-related observables and corresponding data is not likely to be resolved by the introduction of additional complexity in the mantle flow theory, such as lateral viscosity variations or lateral heterogeneity in major element mantle composition.

The main reason for the substantial misfit of the tomography-based geodynamic predictions is most likely due to a basic limitation inherent in published tomography models derived solely from global seismic data. This limitation arises because the data sets employed in the derivation of these tomography models do not provide unique constraints on lateral heterogeneity throughout the mantle. It has long been understood that the amplitude of seismically inferred heterogeneity is poorly constrained over a large depth interval in the mid and lower mantle and that direct comparisons between different models published in the literature yield ambiguous information (e.g. Masters et al. 2000, de Wit *et al.* 2012). This ambiguity or non-uniqueness is the direct consequence of the existence of a large null space in the seismic data constraints on lateral heterogeneity in the mantle (e.g. de Wit *et al.* 2012).

The relatively weak global seismic constraints in the mid-mantle region were explicitly shown in Forte *et al.* (1994) where it was also demonstrated that the seismic sampling of heterogeneity in the southern hemisphere portion of the deep mantle was weak. This limited sampling of mantle heterogeneity in the southern hemisphere is especially marked in tomography models derived predominantly from direct arriving seismic waves (e.g. Li *et al.* 2008). A clear illustration of the outstanding difficulty in resolving mantle heterogeneity in the depth interval 600 to 2000 km is provided by the comparison of recently published tomography models presented by Ritsema et al. (2011; see Figs. 1 & 11 in their paper).

The mid-mantle region is of great importance in geodynamic modeling because the mantle flow excited by buoyancy forces in this region explains a significant fraction of the convection-related surface data such as the long wavelength free-air gravity anomalies (see Fig. 10c,d) and the tectonic plate motions (see Fig. 10b). An accurate resolution of mid-mantle buoyancy forces, over a large depth interval (from about 400 to 2000 km depth), is also crucial for understanding the amplitude and spatial pattern of flow across the entire mantle, with direct implications for the extent to which the flow is vertically stratified (e.g. Forte & Woodward 1997a,b, Forte 2000, Simmons *et al.* 2006).

Currently published seismic tomography models do not represent the full range of models that are compatible with the global seismic data sets (e.g. de Wit *et al.* 2012). Although the tomographic inversions may contain far more data than the parameters used to represent the 3-D models (e.g. Boschi & Dziewonski 2000, Li *et al.* 2008, Kustowski *et* *al.* 2008, Ritsema *et al.* 2011), the data nonetheless provide limited and non-uniform sampling of deep mantle heterogeneity. This limitation arises in large part from the geographically restricted distribution of seismic stations and earthquake sources. Smoothness and amplitude-damping constraints are therefore used to regularize the tomographic inversions in order to stabilize the heterogeneity in those portions of the mantle that are poorly sampled by the seismic data. If these regularization conditions are changed, then so are the tomography models that result from the inversions. Furthermore, the result of any seismic inversion will be dependent on a number of other factors that include: model parameterization, data weighting, seismic wave theory employed in the data interpretation. Changes in any one of these ingredients will result in different 3-D seismic models that will yield different and, in some cases, conflicting geodynamic interpretations of mantle heterogeneity.

Since any individual seismic tomography model is only a single realization out of the infinitely large space of models that are compatible with the global seismic data, it is not sufficient to test mantle-flow hypotheses, or to evaluate the significance of predicted geodynamic surface observables, on the basis of the published tomography models. As initially shown by Forte *et al.* (1994) and subsequently by Simmons et al. (2006, 2007, 2009, 2010), a robust test of different mantle-flow scenarios, and the degree to which mantle heterogeneity is controlled by lateral temperature variations, is to check the plausibility of these scenarios directly against the widest range of available seismic and geodyamic data. This testing can be effectively accomplished by simultaneously inverting all data constraints in a single large-scale inversion. These joint inversions will immediately reveal the extent to which these different constraints on mantle structure, composition, and flow, can be simultaneously and successfully reconciled by at least one model of 3-D mantle heterogeneity.

4.3 Simultaneous inversions of seismic, geodynamic and mineral physical data

In essence, a joint tomographic inversion of multiple and distinct classes of geophysical data involves the search for (a) 3-D mantle heterogeneity model(s) that can simultaneously satisfy all these data. This was the approach developed by Forte *et al.* (1994) and Forte & Woodward (1997), who implemented a method for jointly inverting global seismic and geodynamic data. The key feature of these joint inversions, that sets them apart from classical tomography inversions of seismic data alone, is the inclusion of constraints directly relating to buoyancy-induced mantle flow, in particular the relationship between this flow and convection-related surface data. These joint inversions are thus able to directly test to what extent the seismic constraints on mantle heterogeneity are compatible with entirely independent geodynamic constraints on the same structure (e.g. Simmons *et al.* 2006). This joint inversion procedure also allows a direct test of the degree to which mantle heterogeneity is controlled by the thermal anomalies maintained by the mantle convection process (e.g. Forte & Woodward 1997, Simmons *et al.* 2009)

4.3.1 Geodynamic and seismic data sets

The principal geodynamic surface observables that can be used to effectively constrain the 3-D structure and dynamics of the convecting mantle were described previously in section 2.1 and they include the global free-air gravity anomalies, the crust-corrected dynamic topography, the present-day tectonic plate velocities, and the excess 'flattening' or ellipticity of the CMB (all fields are shown above, in Fig. 2).

The theoretical, linear-integral relationship between the spherical harmonic coefficients of any one of these observable geodynamic fields and the corresponding harmonic coefficients of the internal density perturbations in the mantle is provided by geodynamic kernel or 'response' functions, as shown above in equation (104). In the recent series of joint seismic-geodynamic tomography inversions carried out by Simmons *et al.* (2007, 2009, 2010), the geodynamic response functions have been calculated on the basis of the radial viscosity profile, shown in Fig. 9, derived from the simultaneous inversions of GIA and convection data by Mitrovica & Forte (2004). The depth-dependent geodynamic kernel functions that are derived on the basis of this geodynamically inferred viscosity profile are presented in Fig. 10. These kernels will be employed in the joint tomography inversions described below.

The seismic data employed in the joint seismic-geodynamic tomography inversions consist of over 46,000 travel times of shear body waves (S waves) that were measured by Simmons (2007) and Grand (2002) through cross-correlation of synthetic waveforms (for 1-D reference Earth model) and observed waveforms. These travel-time measurements were made mainly on teleseismic, horizontally-polarized (SH) shear waves, including reflected multiple phases: Sn and $ScSn$, where n is the number of surface reflections and c denotes a reflection from the CMB. The data set also contains vertically-polarized (SV) teleseismic shear phases: SKS and $SKKS$. An additional, critically important component of the global data set consists of regional, shallow-turning SH phases and surface-reflected multiples that are triplicated due to sharp velocity increases in the upper-mantle transition zone (Grand 1994). These shallow phases provide increased lateral resolution of strongly-varying upper-mantle structures that teleseismic body waves do not provide. A thorough discussion of the seismic data measurement and modelling procedure is provided by Simmons (2007), with a detailed summary to be found in Simmons *et al.* (2006).

4.3.2 Joint inversion procedure

The joint inversion of the seismic and geodynamic data sets described in the preceding section is carried out by first parameterising the heterogeneity in the mantle in terms of individual blocks, in each of which the heterogeneity is assumed to be constant. This block representation of 3-D mantle structure is accomplished by dividing the mantle into 22 layers, from the surface down to the CMB, where the layer thickness varies between 75 km (in the upper mantle) to 150 km (in the lower mantle). The lowermost mantle layer (layer 22), representing the heterogeneity in the seismic D" region, has a thickness of 240 km. Each layer is then divided into rectangular voxels whose sides have a horizontal dimension of 275 km. This parameterisation of mantle structure yields nearly 99,000 blocks.

The suite of convection-related geodynamic observables (Fig. 2) are combined with the global seismic data into a single linear system, yielding an inverse problem that simultaneously constrains seismic velocity and density heterogeneity as follows:

$$\begin{pmatrix} \mathbf{L} \\ \lambda_G \, \mathbf{G}(R_{\rho/s}) \\ \lambda_{CMB} \, c(R_{\rho/s}) \\ \lambda_R \, \mathbf{R} \end{pmatrix} \mathbf{m} = \begin{pmatrix} \mathbf{r} \\ \lambda_G \, \mathbf{s} \\ \lambda_{CMB} \, e \\ \mathbf{0} \end{pmatrix} \tag{109}$$

in which the heterogeneity in the mantle, denoted here by the column vector **m** with ∼99,000 components, is expressed in terms of seismic slowness perturbations. The seismic constraints on mantle heterogeneity are incorporated via **L**, a sparse matrix of ray path lengths (seismic sensitivities), and **r**, the shear wave travel-time residuals. **L** has a row dimension of ∼46,000, corresponding to the total number of seismic travel times. The geodynamic constraints are incorporated through **G**, a fully populated matrix containing the viscous flow sensitivity kernels (Fig. 10), where each row corresponds to a specific spherical harmonic component of the free-air gravity, plate divergence and dynamic topography constraints. Since each of these fields is expressed by truncated spherical harmonic expansions up to a maximum harmonic degree $\ell = 16$, there are a total of 861 rows in **G**. Each component of the column vector **s** contains the spherical harmonic coefficients of the geodynamic observables (Figs. 2a–c) normalized by their estimated standard error. The row vector c is the viscous flow response function (i.e. kernel) for the $\ell = 2$ zonal harmonic coefficient describing the excess ellipticity of the CMB (Fig. 2d), contained in the data component e.

The inversion is stabilized through the introduction of a regularization or smoothing operator represented by the matrix **R**, designed to yield a second-order smoothing in both the vertical (layers) and horizontal (blocks) directions (Simmons 2007). The λ terms in equation (109) are scalar quantities that determine the relative weights applied to the seismic, geodynamic and smoothing constraints.

Since the geodynamic data are strictly related to density perturbations rather than the seismic heterogeneity (see section 3.2), we must also incorporate a relationship between perturbations of density (ρ) and shear wave velocity (V_S). In equation (109), the density-velocity relationship is identified by $R_{\rho/s}$ which is defined by:

$$R_{\rho/s} = \frac{\mathrm{dln}\rho}{\mathrm{dln}V_S} \tag{110}$$

The selection of appropriate density-velocity scaling is clearly important in ensuring that the independent constraints on mantle heterogeneity provided by the seismic and geodynamic data can be fully reconciled. This reconciliation also bears on the extent to which this heterogeneity is dominated, or not, by the thermal structure maintained by the convecting mantle. These considerations are discussed in the next section.

Finally, the joint linear system in (109) is inverted using an iterative LSQR inversion technique (Paige & Saunders 1982). Special attention is required to ensure successful convergence of the iterations, owing to the mixed (i.e. sparse and fully populated) character of the joint matrix on the left-hand side of expression (109). A modified LSQR scheme was therefore implemented by Simmons (2007). This modified scheme also permits an independent specification of the density-velocity scaling in (110), such that it is fully decoupled from the geodynamic kernels in expression (105). In a further modification introduced by Simmons et al. (2010), the LSQR iterations are embedded in a very fast simulated annealing process that ensures maximum consistency between the seismic and geodynamic constraints on 3-D mantle structure.

4.4 Relating mantle density and seismic velocity anomalies

The principal hypothesis employed in previous joint inversions of global seismic and geodynamic data sets (Forte *et al.* 1994, Forte & Woodward 1997b, Simmons *et al.* 2007, 2009, 2010) is that the lateral temperature variations maintained by the process of thermal convection are the dominant source of lateral heterogeneity in the mantle. The most notable and generally accepted exception to this thermal dominance is the heterogeneity associated with the so-called 'tectosphere' (e.g. Jordan 1981), namely the deep lithospheric roots below ancient continental shields, wherein geodynamic data consistently show that significant compositional heterogeneity occurs and that it almost completely offsets the negative buoyancy of thermally-generated density anomalies (e.g. Forte & Perry 2000).

The other, more speculative exception concerns the possible dominance of large-scale thermo-chemical 'piles' at the very bottom of the mantle (e.g. McNamara & Zhong 2005), that could be interpreted as the consequence of long-term accumulation of subducted MORB-derived heterogeneity in the seismic D″ layer (e.g. Nakagawa & Tackley 2008). The dominant signature of these chemical piles is less generally accepted because mantle convection modelling has shown that most of the seismic (shear-velocity) heterogeneity in the lower mantle can be explained quantitatively in terms of lateral temperature variations (e.g. Quéré & Forte 2006, Schuberth *et al.* 2009). The possibility that lateral temperature variations maintained by the convecting mantle are the dominant source of 3-D mantle heterogeneity is further supported by recent seismic modelling of mantle structure generated by theoretical convection models (Schuberth *et al.* 2012).

Whether or not compositional heterogeneity is a dominant contributor to 3-D seismically inferred structure in the lithospheric mantle below continents, or in the seismic D″ layer, the main driving force for the convective flow is due to buoyancy forces in the bulk of the mantle that are generated by hot thermal plumes and cold subducting lithosphere. This fundamental fluid-mechanical property of thermal convection, in which the global circulation is mainly driven by the buoyancy associated with thermal anomalies in the interior of the convecting layer (i.e. outside the thermal boundary layers at the top and bottom of the layer), has been understood and established long ago (e.g. Turcotte & Oxburgh 1967, Jarvis & Peltier 1982). It is therefore clear that constraining the amplitude and spatial distribution of thermal anomalies in the bulk of the mantle holds the key to understanding the dynamics of the convecting mantle.

In view of the preceding considerations, the natural starting assumption to employ in the simultaneous tomographic inversion of global seismic and geodynamic data sets is an *a-priori* relationship between seismic shear-velocity and density anomalies, $R_{\rho/s}$, defined in eq. (110) that is expressed in terms of lateral temperature variations and is accordingly quantified in terms of mineral physics data. In the recent series of inversions carried out by Simmons *et al.* (2007,2009,2010), the iterative determination of lateral heterogeneity in the mantle that simultaneously explains seismic and geodynamic data was commenced with a depth-dependent $R_{\rho/s}$ based on the mineral physical results of Cammarano *et al.* (2003) for the upper mantle, and Karato & Karki (2001) for the lower mantle. The optimal 1-D $R_{\rho/s}$ profile that was determined on the basis of these mineral physics studies is represented by the red curve in Fig. 23 (explicit values are provided in Table 3 in Simmons *et al.* 2009).

The joint tomography model derived on the basis of this 1-D density-velocity scaling (red curve in Fig. 23) provides an excellent match to the global seismic data (93% variance reduction) that is essentially identical to a pure seismic model obtained in an inversion without geodynamic constraints (Simmons *et al.* 2009). The joint tomography model also yields a reasonably good match to the global geodynamic constraints: variance reductions of 71%, 96%, and -9% to the free-air gravity anomalies, plate motions (expressed in terms of their horizontal divergence), and crust-corrected dynamic topography, respectively. Although the fits to the gravity anomalies and plate motions are now a great improvement over those obtained with pure seismic tomography models (see Tables 1, 2), it is also clear that a purely thermal interpretation of mantle heterogeneity is not compatible with the dynamic topography data. The principal source of topography misfit in continental regions is the neglect of compositional heterogeneity under ancient cratons, as discussed above, and the non-linear thermal effects associated with temperature-dependent seismic attenuation and partial melting in the shallow mantle, under oceanic (especially mid-ocean ridges) and tectonically-active regions (Simmons *et al.* 2009).

To account for the impact of this upper-mantle compositional heterogeneity under cratons and for non-linear thermal effects, Simmons *et al.* (2009, 2010) introduced an effectively 1.5-D or 'corrected' $R_{\rho/s}$ scaling relationship, defined as follows:

$$R_{\rho/s}^{corrected} = R_{\rho/s}^{1-D} + \kappa\, \delta\ln V_S, \quad \text{where} \;\; \kappa = \left(\frac{\partial R_{\rho/s}}{\partial \ln V_S}\right) \tag{111}$$

The partial derivative κ adjusts or corrects the density-velocity scaling $R_{\rho/s}$ according to the seismic heterogeneity model, $\delta\ln V_S$, obtained in the preceding joint inversion assuming the radially-symmetric $R_{\rho/s}^{1-D}$ model (red curve in Fig. 23).

Simmons *et al.* (2009) conducted a search for optimal values of κ in expression (111) by first identifying three distinct regions in the upper mantle: (1) cratonic lithospheric roots, from 25 to 250 km depth, (2) non-cratonic shallow mantle, from 25 to 250 kim depth, and (3) deep upper mantle, from 250 to 650 km. The values of κ in regions (2) and (3) model the impact of non-linear, purely thermal effects in the upper mantle. The value of κ in region (1) models the impact of the compositional heterogeneity in the cratonic upper-mantle associated with the continental tectosphere. Simmons *et al.* (2010) further refined this regionalization of the shallow mantle by searching for separate κ values in the depth intervals 25–100 and 100–250 km.

The κ values in eq. (111) that provide an optimal reconciliation of the joint seismic and geodynamic (especially dynamic topography) constraints on mantle heterogeneity are summarized in Table 3 in Simmons *et al.* (2009) and in Table 4 in Simmons *et al.* (2010). On the basis of this effectively 1.5-D parameterisation of $R_{\rho/s}$, the variance reductions to the geodynamic data reported by Simmons *et al.* (2009) are 75% for free-air gravity, 98% for horizontal plate divergence, and 41% for dynamic topography. In the joint inversions of Simmons *et al.* (2010) the variance reductions for 1.5-D $R_{\rho/s}$ are 73% for free-air gravity, 97% for horizontal plate divergence, and 32% for dynamic topography. The fit (variance reduction) to the global seismic shear-velocity data is 93% for Simmons *et al.* (2009) and 90% for Simmons *et al.* (2010), where the latter inversion is also constrained to satisfy global P-wave travel-time residuals.

These upper-mantle corrections to the purely 1-D, thermal $R_{\rho/s}$ scaling profile, in particular under continental cratons, yield a substantially improved fit to the dynamic topography but it is still clear that this geodynamic observable is not reproduced as well as the gravity or plate motion data. One possible explanation for this deficiency is that the 1.5-D corrections, as modelled in expression (111), do not accurately reproduce a number of mineral physical effects, that include (1) additional compositional heterogeneity in the cratonic roots not explained in terms of the simplified velocity-based correction; (2) compositional heterogeneity elsewhere in the mantle; (3) partial melting and strongly nonlinear (e.g. seismic-Q dependent) effects that cannot be simply explained by the velocity-based correction.

The more complex, compositional or non-thermal effects not accounted for in the corrected 1.5-D density-velocity scaling in expression (111) can be determined by solving for a fully 3-D $R^{3-D}_{\rho/s}$ scaling relationship between seismic shear velocity and density anomalies. The determination of this completely general density-velocity scaling relationship is an intrinsically nonlinear problem that is based on the linear system in eq. (109) that defines the joint tomography model, and it yields $R^{3-D}_{\rho/s}$ by iteratively solving the following linear system:

$$\begin{pmatrix} \lambda_G\, \mathbf{G}(\mathbf{m}^k) \\ \lambda_{CMB}\, c(\mathbf{m}^k) \\ \lambda_R\, \mathbf{R} \end{pmatrix} \mathbf{R}^{3-D,k+1}_{\rho/\mathbf{s}} = \begin{pmatrix} \lambda_G\, \mathbf{s} \\ \lambda_{CMB}\, e \\ \mathbf{0} \end{pmatrix} \tag{112}$$

where $\mathbf{m}^k$ is the seismic slowness model obtained by joint inversion of system (109) assuming a density-velocity scaling $R^{3-D,k}_{\rho/s}$ obtained in a previous cycle of the iterative process. System (112) is solved via an LSQR inversion, as the original system (109), except that the unknowns are the block representation of the density-velocity scaling $\mathbf{R}^{3-D,k+1}_{\rho/\mathbf{s}}$, for a fixed or given seismic heterogeneity model. At the very outset, for $k = 0$, the iterations are initiated with the corrected 1.5-D scaling defined above in expression (111). The next iterative update of the density-velocity scaling, $R^{3-D,k+1}_{\rho/s}$ is then found by inverting expression (112) and the cycles continue until convergence.

In Simmons *et al.* (2009) only the first $k = 0$ cycle was solved in eq.(109), where the final seismic-shear heterogeneity model was determined on the basis of the 1.5-D corrected density-velocity scaling (111), and the final 3-D scaling was found by solving eq. (112) once. Figure 23a provides a graphical summary of the lateral variability of the 3-D density-velocity scaling $R^{3-D}_{\rho/s}$ relative to the 1-D thermally derived scaling. The large scatter in the top 300 km of the mantle, extending to negative values, represents strong compositional contributions in the sub-cratonic mantle under ancient continental shields. In the upper-mantle transition zone the modes of the histograms representing the lateral variability are displaced relative to the 1-D thermal profile, suggesting that some revision to the depth-dependent mineral physical estimates may be warranted. This may be due to compositional or mineralogical (phase-change dependent) complexity that is not modelled with the thermally-derived background scaling. In the lower mantle, below 2000 km depth, there is also similar evidence of displaced modal values relative to the thermal scaling. Nonetheless, the outstanding characteristic of the laterally variable $R_{\rho/s}^{3-D}$ in the lower mantle is the close grouping of the scaling values around the *a-priori* mineral physical estimates and this demonstrates that a dominantly thermal origin for heterogeneity in the mantle is fully compatible with the combined seismic and geodynamic constraints on the 3-D mantle structure.

Simmons *et al.* (2010) iteratively inverted the combined system of equations (109) and (112) through about 60 cycles to achieve convergence. In each cycle the initial, thermally based 1-D scaling relationship (red curve in Fig. 23a) and the non-linear thermal effects, as modelled by eq. (111), were also corrected and revised. The final, optimal 1-D scaling is shown by the red curve in Fig. 24a and the distribution of laterally variable 3-D scaling at each depth in the model, inferred by solving system (112), is represented by the shaded histograms. In the upper 300 km of the mantle we again note the strongly elongated distribution, extending to negative values, that is needed to characterize the sub-continental mantle under cratons and strongly nonlinear effects (e.g. partial melting) under oceans. In the deeper mantle, the modes of the 3-D scaling distribution are all centred on the background 1-D scaling values (red curve) that has been adjusted relative to the initial thermally-derived 1-D scaling. In the bottom 500 km of the lower mantle the scaling distribution extends to zero values and these characterize the compositional, intrinsically dense interiors of the 'superplume' structures under the African and Pacific plates (see next section).

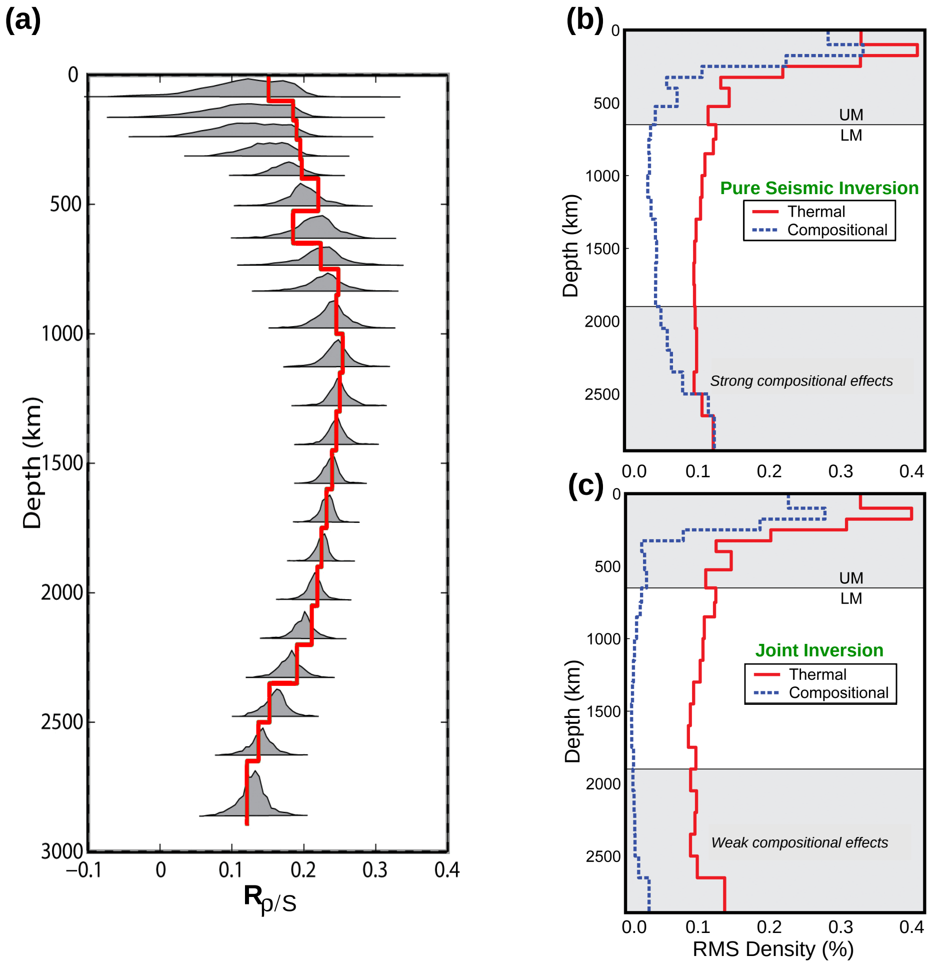

Figure 23: Implications of joint seismic-geodynamic-mineral physical constraints for thermal and non-thermal contributions to mantle heterogeneity. **(a)** The red curve represents the optimal 1-D density-velocity scaling $R_{\rho/s}$ derived from mineral physical data assuming that thermal contributions to heterogeneity are dominant (see text for details). The shaded histograms at each depth represent the normalized distribution of the 3-D density-velocity scaling $R^{3D}_{\rho/s}$ derived from the joint seismic-geodynamic inversions (Simmons *et al.* 2009). **(b)** The red curve represents the thermal contributions to mantle heterogeneity when a seismic tomography model, $TX2008s$ derived by Simmons *et al.* 2009 from seismic data alone, is scaled to density using the 1-D thermal $R_{\rho/s}$ profile shown in (a). The subsequent inversion for a 3-D density-velocity scaling $R^{3D}_{\rho/s}$ that provides a fit to the geodynamic data yields the compositional (non-thermal) contributions to mantle density shown by the blue dashed curve. **(c)** When the seismic and geodynamic constraints are simultaneously inverted, in an iterative inversion that is initiated with the thermal $R_{\rho/s}$ profile shown by the red curve in panel (a), the thermal and compositional contributions to mantle heterogeneity in tomography model $TX2008$ (Simmons *et al.* 2009.) are shown by the red and blue dashed curves, respectively. (Figures adapted and modified from Simmons *et al.* 2009.)

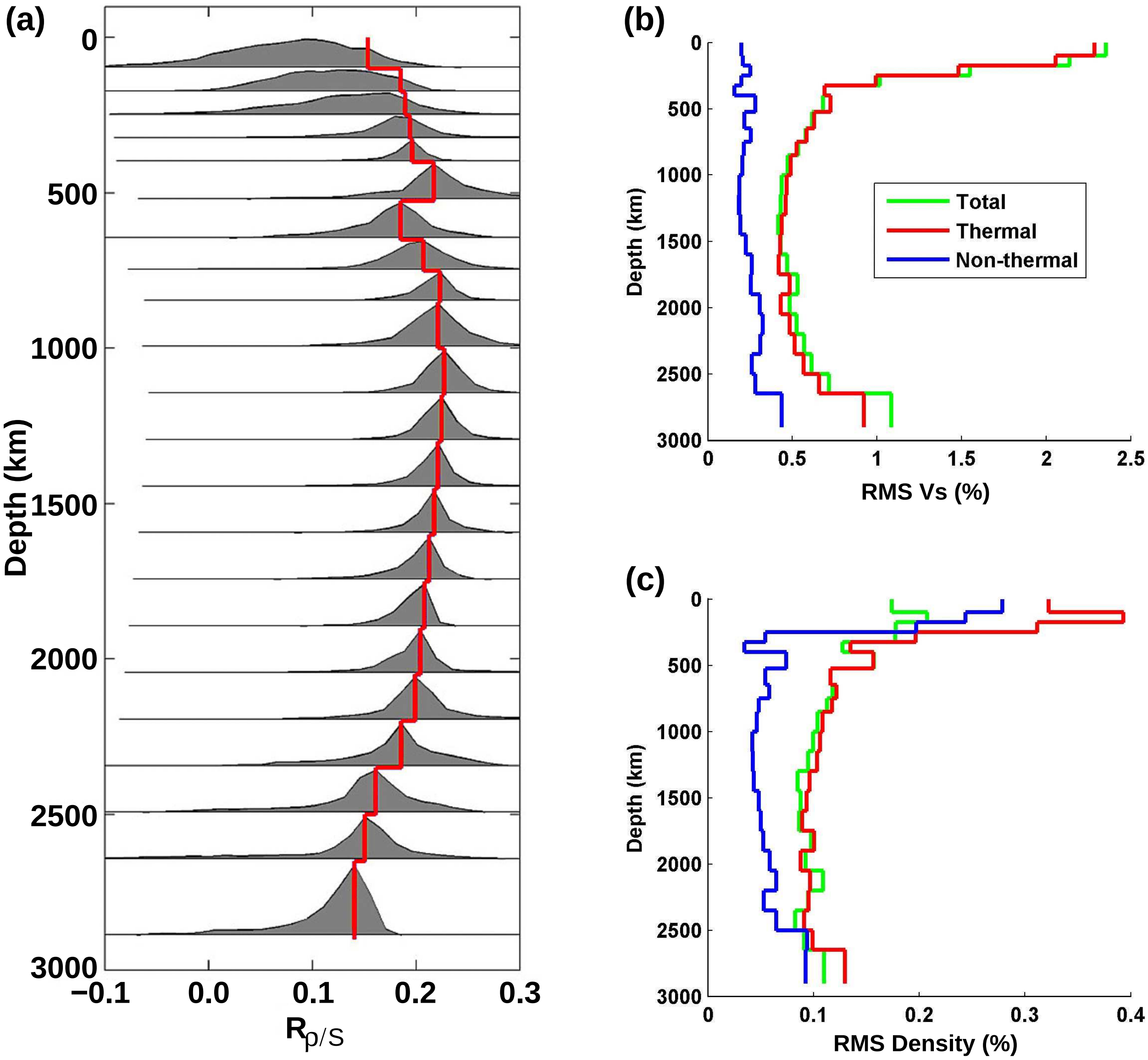

Figure 24: Importance of thermal and compositional heterogeneity from multiparameter joint seismic-geodynamic-mineral physical inversions for Vs, Vp, and ρ anomalies in the mantle. **(a)** The red curve represents the 1-D density-velocity scaling $R_{\rho/s}$ derived from an iterative, simulated annealing optimization that was initiated with the purely thermal scaling profile based on mineral physics data (red curve in Fig. 23a). The grey, filled regions represent the normalized distribution of the 3-D density-velocity scaling $R^{3D}_{\rho/s}$ derived from the joint seismic-geodynamic inversions (Simmons *et al.* 2010). **(b)** The green curve represents the total (rms) amplitude of mantle Vs anomalies in seismic tomography model $GyPSuM$ derived from joint inversion of seismic, geodynamic and mineral physical data (Simmons *et al.* 2010). The red and blue curves represent the thermal and compositional (non-thermal) contributions, respectively. **(c)** As in panel (b), except that mantle density anomalies in model $GyPSuM$ are represented. (Figures adapted and modified from Simmons *et al.* 2010.)

4.5 Joint seismic-geodynamic mapping of density anomalies

The joint seismic-geodynamic tomography inversion for mantle density anomalies is accomplished by solving expression (109), which is characterized by a nonlinear coupling of the seismic shear-velocity heterogeneity $\delta\ln V_S$ and the density-velocity scaling $R_{\rho/s}$. As discussed in the preceding section, the solution is obtained through the iterative inversion of the coupled system of linear equations (109) and (112). The final iteration of this coupled system yields a mapping of the total density perturbation field $\delta\ln\rho^{total}$ that includes both thermal and non-thermal, compositional effects:

$$\delta\ln\rho^{total} = R^{3-D}_{\rho/s}\ \delta\ln V_S^{final} \tag{113}$$

in which $R^{3-D}_{\rho/s}$ and $\delta\ln V_S^{final}$ are the final 3-D density-velocity scaling and seismic-shear heterogeneity, respectively, derived from the iterative inversion process.

The purely thermal contributions to the mantle density anomalies can be estimated using the following expression:

$$\delta\ln\rho^{thermal} = R^{thermal}_{\rho/s}\ \delta\ln V_S^{final} \tag{114}$$

in which the thermal density-velocity scaling $R^{thermal}_{\rho/s}$ is obtained from expression (111), in which the 1-D profile $R^{1-D}_{\rho/s}$ is derived from mineral physical data for thermal effects and the parameterisation of the non-linear thermal effects in the upper mantle is accounted for by the κ term. In the latter case we exclude the use of κ values estimated for the cratonic upper mantle, since these model compositional effects. Explicit values for these thermal terms may be found in Tables 3 and 4 of Simmons *et al.* (2009) and Simmons *et al.* (2010), respectively.

Finally, the compositional or non-thermal contributions to the mantle density anomalies may be simply estimated by subtracting expressions (113) and (114):

$$\delta\ln\rho^{non-thermal} = \delta\ln\rho^{total} - \delta\ln\rho^{thermal} \tag{115}$$

This definition of non-thermal contributions can also include effects not directly modelled by the assumed thermal scaling in (111), such as possible higher-order non-linear behaviour due to Q-dependence of seismic shear velocity or partial melting.

Current efforts to constrain the magnitude of compositional contributions to mantle density perturbations, in particular the hypothesized lower-mantle chemical 'piles', are important to understanding the thermal and chemical evolution of the convecting mantle (e.g. Nackagawa *et al.* 2010, Davies *et al.* 2012). These analyses must contend with the likelihood of bias introduced by *a-posteriori* conversion of published tomography models, derived from seismic data alone, into equivalent maps of density anomalies. The possibility of this bias was demonstrated in an experiment conducted by Simmons *et al.* (2009), in which they determined a 3-D density-velocity scaling by inverting system (112), in which the seismic heterogeneity model $\mathbf{m}$ is independently derived from a classical inversion involving seismic data alone. The amplitude of the thermal and compositional density anomalies, calculated according to expressions (114) and (115), are shown in Fig. **??**b where we note this *a-posteriori* analysis of a pure seismic heterogeneity model yields the inference of strong compositional heterogeneity in the lower mantle. In contrast, when the density and seismic shear-velocity anomalies are simultaneously determined in a joint inversion that is initially constrained by mineral physical data (red curve in Fig. 23a), the inferred amplitude of compositional heterogeneity in the lower mantle is much weaker (Fig. 23c).

The most recent effort at constraining the thermal and compositional heterogeneity in the mantle, via simultaneous inversion of seismic, geodynamic and mineral physics data was carried out by Simmons *et al.* (2010). In contrast to the previous joint inversions by Simmons *et al.* (2007, 2009), the seismic shear-velocity data were augmented to include a very large set of global P-wave travel times. The joint inversions were therefore designed to simultaneously map lateral variations in ρ, V_S and V_P throughout the mantle. The relative amplitudes of thermal and compositional contributions to the seismic shear-velocity heterogeneity, where the former is determined in an inversion based on the thermally dependent version of the 1.5-D scaling in eq. (111), are summarized in Fig. (Fig. 24b). A substantially different illustration of the relative importance of thermal and compositional contributions is provided by the amplitudes of the density anomalies shown in (Fig. refdens-veloc-2010c), where we observe that non-thermal and thermal contributions (obtained via eqs. 114 & 115) are of similar magnitude in the top 250 and bottom 500 km of the mantle. This constrasts with the clear dominance of thermal effects at all depths in the case of the seismic shear-velocity heterogeneity. The latter result lends further support to the conclusion reached by Forte & Mitrovica (2001) that seismic shear-velocity anomalies, in particular in the deep lower mantle, provide a reliable mapping of the lateral temperature variations maintained by the mantle convective circulation.

The field of total density anomalies in the mantle as defined by expression (113) and obtained in the joint inversions by Simmons *et al.* (2009) is referred to as model $TX2008$. The corresponding model of total density anomalies obtained in the joint inversions by Simmons *et al.* (2010) is referred to as model $GyPSuM$. Model $TX2007$ is an earlier density heterogeneity model, essentially a precursor to $TX2008$ but with weaker smoothing constraints, that was obtained by Simmons *et al.* (2007). Maps of all three models at selected depths in the upper and lower mantle are provided in Fig. 25 and Fig. 26, respectively.

The amplitudes of the upper-mantle density anomalies (Fig. 25) generally do not exceed a value of about 0.5%. For example, under the western Pacific subduction zone, in the transition zone (depth interval 525-650 km), the maximum amplitude of the positive density anomalies that correspond to subducted slabs is +0.5%. We can understand the plausibility of this value by considering an idealized descending lithospheric slab with average temperature contrast relative to adiabatic mantle of about 700° and having an average width of about 100 km. If we further assume that this signal is 'smoothed' over a horizontal distance of 275 km, corresponding to the lateral dimension of a single block in the tomography model, then we predict a relative density perturbation +0.5% for this block, assuming a thermal expansion coefficient of 2×10^{-5} K^{-1}.

These density anomalies in the upper-mantle transition zone provide important driving forces for the mantle-wide convective flow, especially for surface plate motions (see the divergence kernels in Fig. 10b). Notice, however, that there are other important sources of upper-mantle buoyancy that are not related to subducted slab heterogeneity, in particular the plume-related anomalies under the Eastern Pacific oceanic ridge system that will also provide a strong contribution to the mantle convective flow and the associated surface plate motions.

The maps of lower-mantle density anomalies (Fig. 26) reveal long-wavelength structures that have been long resolved in numerous purely seismic tomography models (e.g. Dziewonki 1984, Grand 2002, Romanowicz 2003). The most noteworthy features in these density maps are the elongated structures of positive buoyancy located deep under the Eastern Pacific oceanic ridge system, as found in the upper mantle (Fig. 25), especially under the East Pacific Rise. The implications of this buoyancy will become evident in the discussion of mantle flow presented in the next section. We also note, especially in the $GyPSuM$ model, the occurrence of compositionally dense material in the core of the hot superplume structures under the central Pacific Ocean and under Southern Africa. The effect of this negative, compositional buoyancy in the central Pacific is to further enhance the importance of the flow driven by the buoyant hot material associated with plumes under the Eastern and Southwestern Pacific Ocean.

Table 5: Global data fits by joint seismic-geodynamic tomography models

Tomography Model	Data Fits[a]				
	Seismic (%)	Gravity[b] (%)	Divergence[b] (%)	Topography[b] (%)	CMB[b] (metres)
$TX2007$	95	89 (69)	94 (92)	73 (62)	400
$TX2008$	93	91 (68)	99 (99)	80 (66)	400
$GyPSuM$	93	88 (65)	99 (99)	72 (60)	400

[a] With the exception of the excess CMB ellipticity (last column), all data fits are expressed as percent (%) variance reductions.

[b] The geodynamic data include the global free-air gravity anomalies, horizontal divergence of present-day plate motions, the crust-corrected dynamic topography and the excess CMB ellipticity ('observed' value = 400 m), all shown in Fig. 2. With the exception of the CMB, the fits cited are for flow predictions and geodynamic observables truncated at spherical harmonic degree $\ell = 16$. The values in parentheses () are fits for predictions and observables truncated at spherical harmonic degree $\ell = 32$. In all cases the geodynamic predictions are obtained on the basis of a flow calculation using the 'V1' viscosity profile in Fig. 9.

The excellent fits to the combined seismic and geodynamic data sets provided by the three joint tomography models is summarized in Table 5. Even a cursory comparison of these fits with those obtained on the basis of previously published, purely seismic tomography models (see Tables 1–4), reveals the power of the joint inversion method in achieving a very satisfactory reconciliation of the independent constraints on mantle heterogeneity provided by the global seismic and geodynamic data sets. It is of course important to underline that this reconciliation is accomplished in the context of a dominantly thermal origin for the lateral heterogeneity in the mantle.

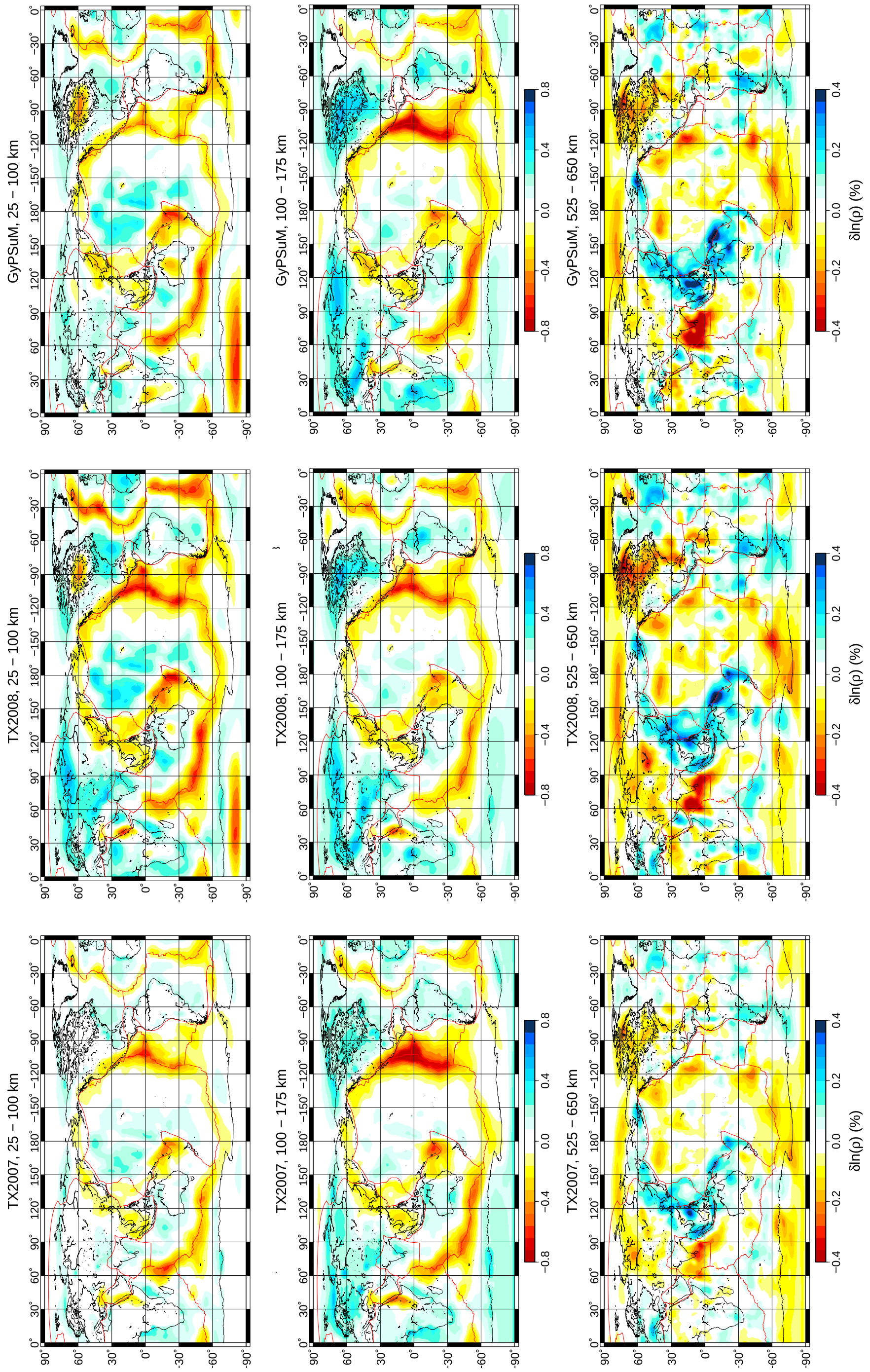
TX2007, 25 – 100 km
TX2008, 25 – 100 km
GyPSuM, 25 – 100 km
TX2007, 100 – 175 km
TX2008, 100 – 175 km
GyPSuM, 100 – 175 km
TX2007, 525 – 650 km
TX2008, 525 – 650 km
GyPSuM, 525 – 650 km
0° 30° 60° 90° 120° 150° 180° -150° -120° -90° -60° -30° 0°
90° 60° 30° 0° -30° -60° -90°
−0.8 −0.4 0.0 0.4 0.8
−0.4 −0.2 0.0 0.2 0.4
δln(ρ) (%)

Figure 25: Global maps of lateral density variations in the upper mantle. *(Left column).* Model $TX2007$ derived by Simmons *et al.* (2007). *(Middle column).* Model $TX2008$ derived by Simmons *et al.* (2009). *(Right column).* Model $GyPSuM$ derived by Simmons *et al.* (2010). A different colour scale is used for the depth interval 25-175 km than for 525-625 km. The density anomalies are represented by a spherical harmonic expansion, as in equation (1), truncated at degree $\ell = 72$, or a minimum half wavelength of ~275 km equivalent to the lateral block dimension employed in the tomography inversions.

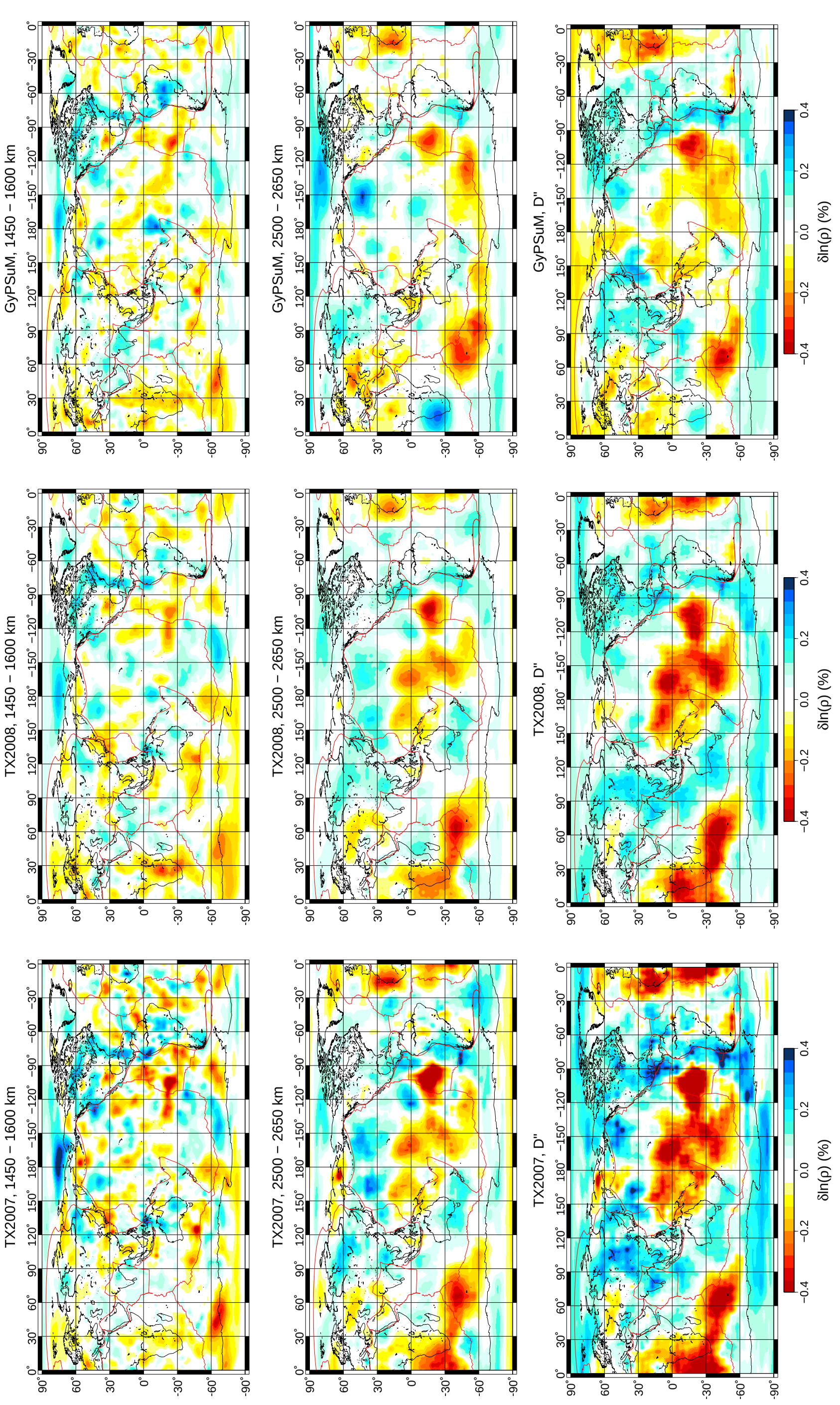
TX2007, 1450 − 1600 km
TX2008, 1450 − 1600 km
GyPSuM, 1450 − 1600 km
TX2007, 2500 − 2650 km
TX2008, 2500 − 2650 km
GyPSuM, 2500 − 2650 km
TX2007, D''
TX2008, D''
GyPSuM, D''
δln(ρ) (%)
−0.4
−0.2
0.0
0.2
0.4

Figure 26: Global maps of lateral density variations in the lower mantle. *(Left column).* Model $TX2007$ derived by Simmons *et al.* (2007). *(Middle column).* Model $TX2008$ derived by Simmons *et al.* (2009). *(Right column).* Model $GyPSuM$ derived by Simmons *et al.* (2010). As in Fig. 25, the density anomalies are represented by a spherical harmonic expansion truncated at degree $\ell = 72$.

4.6 Implications for global mantle flow

In developing a successful or 'realistic' model of the present-day thermal convective flow in the mantle it is clearly essential that such a model be capable of providing a satisfactory fit to a diverse suite of surface geodynamic observables, as summarized in Table 5 in the preceding section. This necessary condition follows from the recognition that these observables provide robust constraints on density anomalies across the entire mantle, as shown by the geodynamic kernel functions in Fig. 10. The density anomalies that are constrained by these surface observables are, of course, those which also drive the mantle-wide convective circulation.

The relationship between the mantle convective flow and the internal density anomalies can be explicitly quantified in terms of Green functions that characterize the viscous-flow impulse response of the mantle to point density loads (e.g. Parsons & Daly 1983, Forte & Peltier 1987). The derivation of these Green functions for a compressible, self-gravitating mantle with a complex depth-dependence of density and viscosity was presented in section 2.3. We recall from section 2.3.7, in particular expression (95), that the viscous-flow Green functions relate the 3-D mantle flow field to the internal density perturbations via the following convolution integral:

$$U^{\alpha m}_{\ell}(r) = \int_b^a \mathbf{U}^{\alpha}_{\ell}(r, r')\rho^m_{\ell}(r')\, dr' \, , \tag{116}$$

in which $\rho^m_\ell(r')$ is the spherical harmonic coefficient of the mantle density perturbations, $U^{\alpha m}_{\ell}(r)$ is the generalised spherical harmonic coefficient of the velocity field (see eq. 31), and $\mathbf{U}^{\alpha}_{\ell}(r, r')$ is the corresponding Green function.

The Green functions for the vertical component of flow, calculated on the basis of the 'V2' viscosity profile (Fig. 9), are plotted at four target depths in Fig. 27. The average amplitudes of the Green functions for depths < 1000 km are significantly greater than those for the deeper mantle, owing the significantly lower viscosity in upper mantle. At shallow depths, in the low-viscosity asthenosphere region, flows characterized by shorter horizontal wavelengths (in the degree range $\ell > 8$) are excited most efficiently. This characteristic persists to bottom of the upper-mantle, in the transition zone, although it is now the intermediate wavelengths (in the range $8 < \ell < 64$) that are more strongly excited. The shorter horizontal wavelengths are more sensitive to the higher-magnitude local viscosity in the transition zone and hence are more attenuated. In the lower mantle a completely different situation arises, where the longest horizontal wavelengths of flow (in the range $\ell < 8$) are excited with the greatest amplitude, because of the very high viscosity in the deep mantle. It is especially noteworthy that these long-wavelength components of the lower-mantle flow field are providing an integrated response to density anomalies across the entire depth of the mantle.

It is instructive to note, as a general rule, that at sufficiently short horizontal wavelengths, for $\ell \gg 1$, the vertical flow at a given target radius r_o in the mantle is sensitive to density anomalies located in a vertical interval $r_o - \Delta r/2 < r_o < r_o + \Delta r/2$, where $\Delta r \approx 2\pi r_o/\ell$. Hence even at shallow depths, for example in the low-viscosity asthenosphere where intermediate to shorter wavelengths are excited strongly, the flow is sensitive to the integral of density anomalies across the entire upper mantle (see top-left panel in Fig. 27). It is therefore difficult to make robust inferences of shallow mantle flow, for example under continents, by only considering 3-D images recovered from regional tomography studies that describe lateral heterogeneity in the top 200 or 300 km of the mantle (e.g. Levander *et al.* 2011), especially when the tomography models are derived from short-period seismic surface waves or ambient noise techniques (e.g. Shen *et al.* 2013). Shallow flow driven by sources of buoyancy located in the transition zone region of the upper mantle will be important and their contribution must be included in flow modelling.

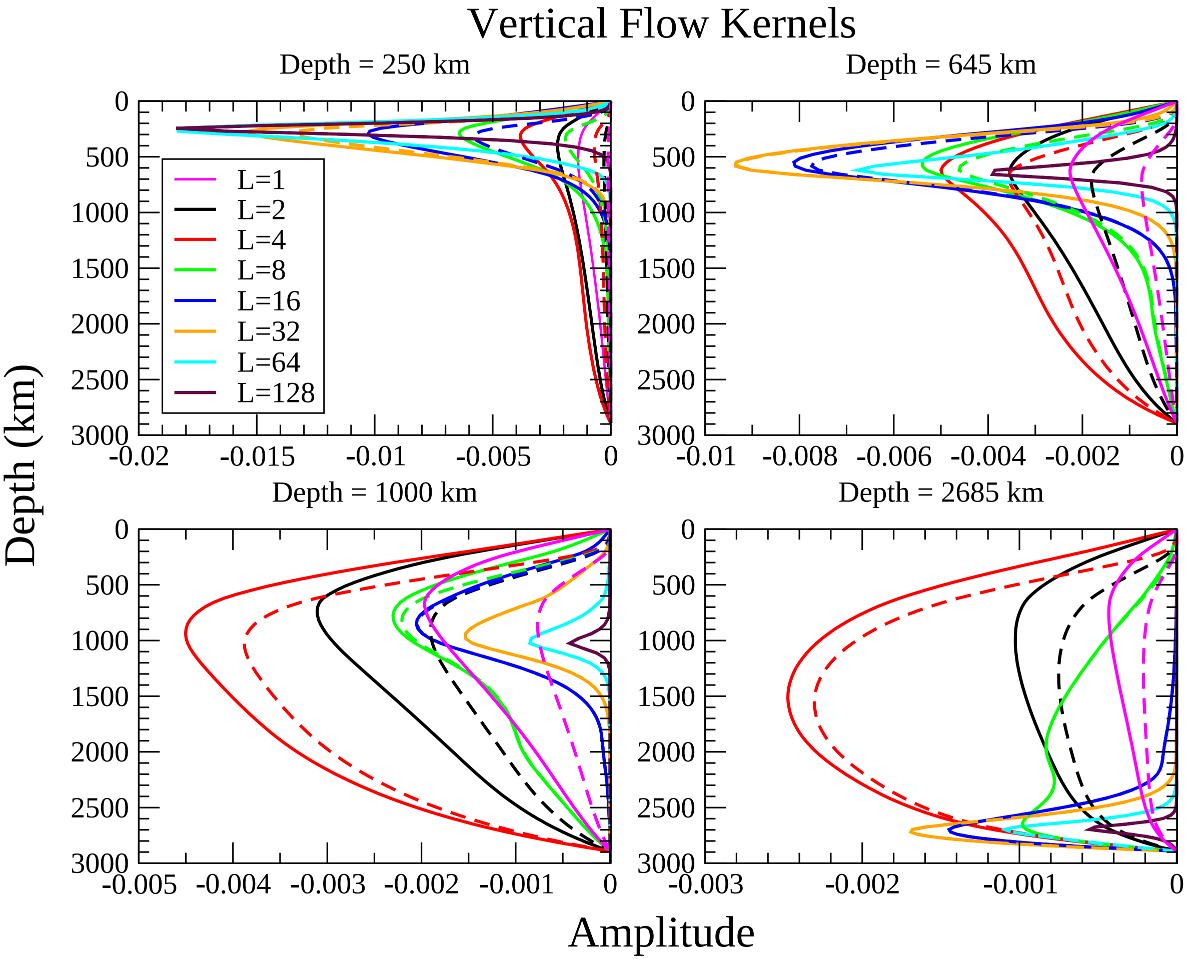

Figure 27: Green functions $U_\ell^0(r, r')$ for the vertical component, $U_\ell^{0m}(r)$ (see eq. 31), of the mantle flow field. The Green functions are plotted for four different depths in the mantle: (*top-left*) 250 km, in the mantle asthenosphere, (*top-right*) 645 km, at the bottom of the transition zone, (*bottom-left*) 1000 km, in the top half of the lower mantle, (*bottom-right*) 2685 km, at the top of the seismic D″ layer. The kernels are colour-coded according to the spherical harmonic degree of the flow (legend in the top-left panel). The solid lines represent kernels calculated for a free-slip surface boundary and the dashed lines represent kernels for a no-slip (rigid) surface. The kernels are calculated for the 'V2' mantle viscosity profile (Fig. 9) and for the depth-dependent density given by PREM (Fig. 5).

There are two distinct contributions to the horizontal component of mantle flow, the poloidal contribution ${U^{P}}^{m}_{\ell}(r)$ (defined in eq. 37) that is directly driven by mantle density anomalies according to expression (116), and the independent toroidal contribution ${U^{T}}^{m}_{\ell}(r)$ (defined in eq. 37) that is viscously coupled to the radial vorticity of the tectonic plates, whose motions are generated by the buoyancy driven (poloidal) mantle flow (see eq. 94).

The Green functions for horizontal poloidal flow are plotted in Fig. 28 at the same four target depths as for the vertical flow kernels (Fig. 27). The horizontal flow kernels provide a sensitivity to mantle density anomalies that is markedly different from that of the vertical flow. Indeed it is quite clear that at sufficiently short horizontal wavelengths ($\ell > 8$), the horizontal flow kernels are effectively a radial derivative of the vertical flow kernels. This is expected on theoretical grounds (see eq. 119 in Forte & Peltier 1994). Horizontal flow will therefore be particularly sensitive to depth-dependent changes in the distribution of the mantle density anomalies. In the deep interior of the mantle (e.g. at depths of 645 and 1000 km – see Fig. 28), this vertical sensitivity is also apparent at the longest horizontal wavelengths. In the deep lower mantle we again note, owing to the high viscosity of this region, that only the longest horizontal wavelengths (corresponding to $\ell < 8$) of horizontal flow are efficiently excited and they will provide an integrated response to density anomalies across the entire mantle.

The viscous excitation function for toroidal flow is defined by the following extension to eq. (94):

$$ {U^{T}}^{m}_{l}(r) = F^{T}_{\ell}(r) \frac{\imath\, r_a}{\Omega_1} \left(\hat{\boldsymbol{r}} \cdot \boldsymbol{\nabla} \times \mathbf{v} \right)^{m}_{l} \ , \tag{117} $$

where $F^{T}_{\ell}(r)$ is the depth-dependent toroidal-flow response, determined by the downward propagation of the surface vector defined in eq. (74). This downward propagation is described above in section 2.3.6. A plot of the toroidal-flow excitation function $F^{T}_{\ell}(r)$ is presented in Fig. 29. We note here that excitation of toroidal flow, through viscous coupling with the overlying plate motions, is most efficient in the upper mantle and very little extends into the lower mantle. The only apparent exception is the degree-1 toroidal flow, which has significant amplitude at all depths in the mantle. This is, however, an entirely degenerate solution of the toroidal-flow equations (see eq. 36) when viscosity varies only with depth and has no dynamical significance.

The numerical solution for the $\ell = 1$ excitation function presented in Fig. 29 is mathematically identical to the simple function $F^{T}_{\ell=1}(r) = (r/r_a)$ and is hence is independent of the mantle viscosity structure. It is possible to show (e.g. Forte & Peltier 1994) that this function yields a horizontal flow in the mantle that is everywhere mathematically identical to a simple rigid-body rotation of the entire mantle:

$$ \mathbf{u}^{toroidal}_{\ell=1}(r) = \boldsymbol{\Omega} \times \boldsymbol{r} \tag{118} $$

where the angular rotation $\boldsymbol{\Omega}$ is identical to the global net rotation of the tectonic plate motions. This rigid-body rotation is associated with a state of zero stress in the mantle and therefore it cannot be generated in a dynamically consistent treatment where the viscosity varies with depth only. Lateral viscosity variations are required to generate degree-1 toroidal flow (e.g. Forte & Peltier 1994). The degree-1 degeneracy that arises with radially symmetric viscosity is eliminated by calculating flow in reference frame with zero net rotation. The latter is dynamically equivalent to the rotating frame described by expression (118) because all flow-induced stresses in the mantle are identical in the two reference frames. For this reason all mantle flow predictions presented below (Figs. 30, 31) employ a global no-net-rotation frame for the mantle.

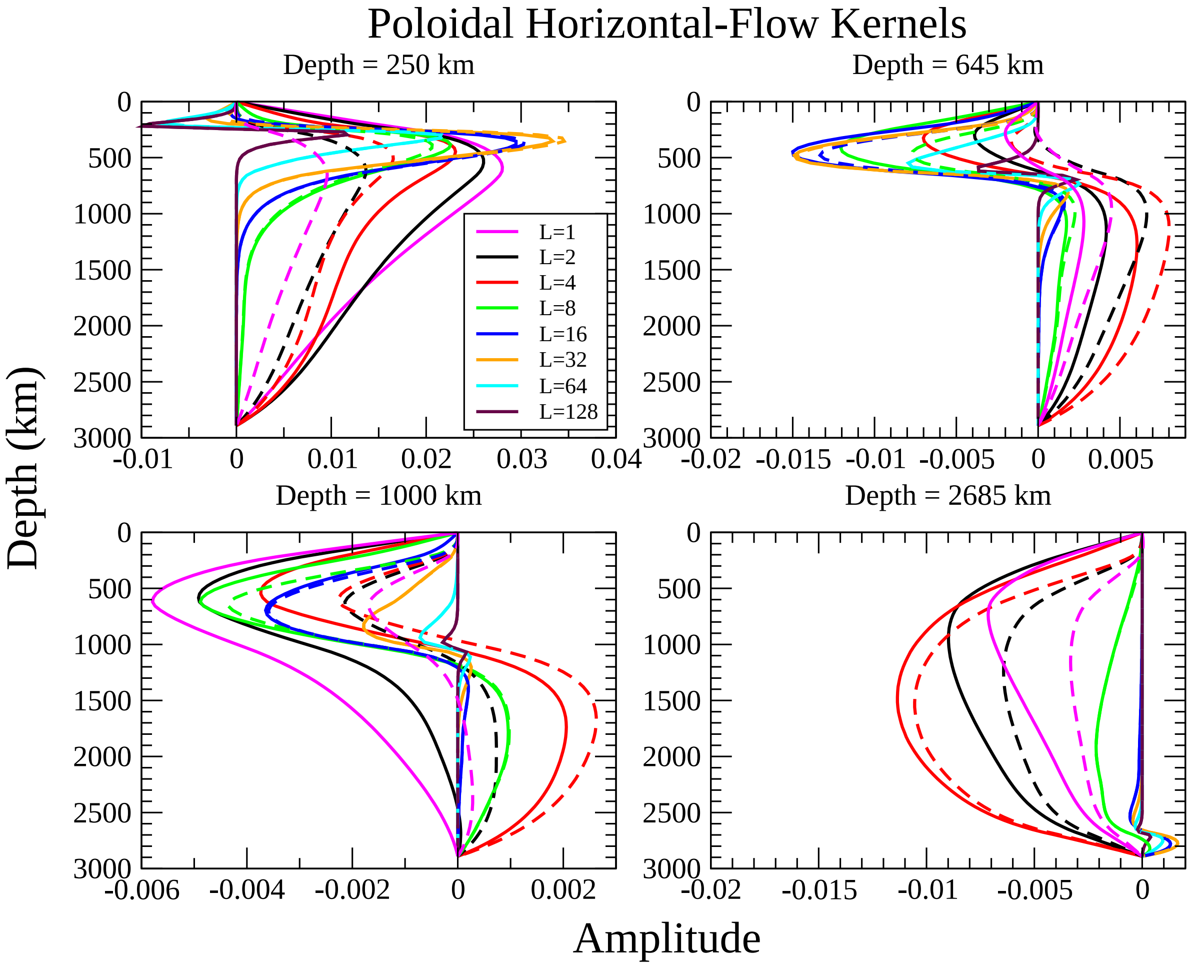

Figure 28: Green functions $U_\ell^P(r, r')$ for the horizontal component, $U^{P}{}_\ell^m(r)$ (see eq. 37), of the poloidal mantle flow field. All calculations and details are as in Fig. 27.

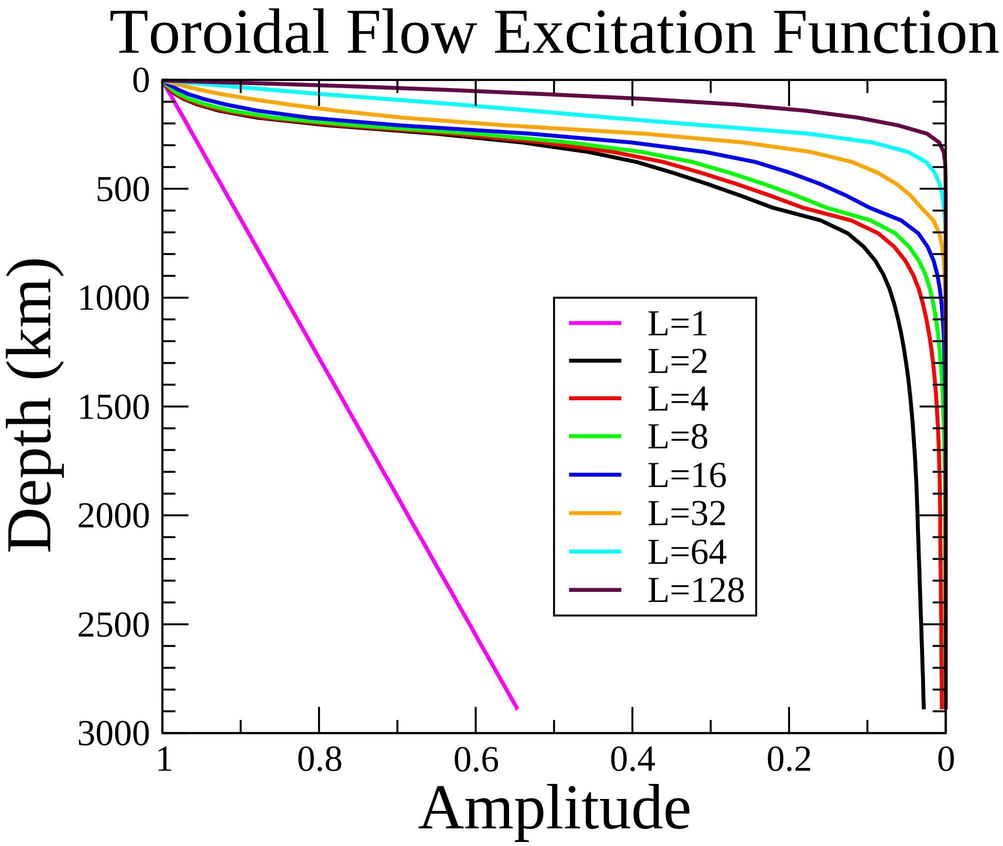

Figure 29: Excitation functions $F^T_\ell(r)$ for the horizontal component, ${U^T}^m_\ell(r)$ (see eq. 37), of the toroidal mantle flow field. These excitation functions are calculated for the 'V2' mantle viscosity profile (Fig. 9).

Given the excellent geodynamic fits provided by the joint seismic-geodynamic tomography density models, it is of interest to explore the implications of these models for the global convective flow in the mantle. Figs. 30 & 31 show the predicted convective flow at selected depths in the upper and lower mantle, respectively, obtained on the basis of the $TX2008$ and $GyPSuM$ tomography models. The predicted flow at the top of the lithosphere (top row in Fig. 30) provides an excellent match to the (No-Net-Rotation) NUVEL-1 plate velocity field. This is not unexpected since the joint seismic-geodynamic tomography models were inverted with the divergence of the surface plate motions as a constraint (see Table 5). This prediction mainly serves as an explicit demonstration that the density anomalies in the mantle, integrated with the kernels in Figs. 28,29, have been properly resolved by the tomographic inversion procedure.

In the low-viscosity asthenosphere (middle row in Fig. 30), the large-scale pattern of vertical flow is similar to that in the lithosphere, with upwellings and downwellings under the oceanic ridges and subduction zones, respectively. The pattern of horizontal flow, however, deviates significantly from the lithospheric pattern in many locations and can even be anti-correlated, for example under the Western U.S. There is an additional complexity that is represented by the striking, 'mottled' pattern of shorter wavelength upwellings that is ubiquitous. For example, local plume-related upwelling under

Hawaii, Iceland and the numerous volcanic swells of the African plate (e.g. Azores, Canaries, Cape Verde, Hoggar Massif, East African Rift) are clearly evident. In addition, and less generally recognized, are the upwellings under the interiors of the Eurasian and North American plates and the many local upwellings under the Pacific plate. The quasi-linear pattern of upwellings that extends from the Caribbean to Eastern Canada, along the Atlantic Coast of the U.S., and the clearly resolved upwelling under Bermuda, have achieved considerable prominence in recent work that reveals the importance of these upwellings in driving variable topographic uplift of the U.S. East Coast (Rowley *et al.* 2013).

The predicted flow at the bottom of the upper mantle, in the transition zone region (bottom row in Fig. 30), strongly reinforces the pattern of (plume-related) upwelling and (subducted-slab) downwelling found at shallower depths. Focussed plume-related upwellings under Hawaii, Iceland, Crozet, Kerguelen, Caroline, Canary, Cape Verde Islands are all evident. A deep, linear upwelling under the southern margin of the Indian plate contributes to strong north-northeast directed horizontal flow (see flow at 250 km) that clearly represents a source of strong (viscous shear) forcing for the continued motion and collision of the overlying Indian plate. Most striking, however, is the intensified pattern of deep upwelling under the East Pacific ridge system. The deep-seated, buoyancy-driven flow under the East Pacific Rise (EPR), in particular, is found to be an important driver of the Pacific plate motion, such that time-dependent changes in this buoyancy have been directly linked to the geologically recent deceleration of this plate (Forte *et al.* 2009).

The lower mantle is a region of very high viscosity (see Fig. 9) and, as anticipated on the basis of the flow kernels (Figs. 27, 28), the predicted flow (see Fig. 31) is dominated by the longest horizontal wavelengths. The long-lived history of lithospheric subduction under the Northern and Southwestern Pacific is clearly imprinted in the pattern of lower-mantle downwellings, as is the history of subduction under the Tethys and Farallon convergence zones. The predicted flow field in the deep lower mantle (middle and bottom row of Fig. 31) reveals several distinct centres of active, buoyancy-driven upwellings, or 'superplumes'. The most prominent superplumes in the model $TX2008$ prediction are located under Southern Africa and Southwest Indian Ocean, under the Southwestern Pacific (under the Caroline Islands), and under the EPR (centred under Easter Island). Additional prominent, but lower-amplitude upwellings are located under western Siberia, under Cape Verde, and under the Southwestern U.S. The upwelling under Siberia is of particular interest because it is correlated with the proposed location of the eruption of the Siberian traps (Smirnov & Tarduno 2010). As shown by the mantle convection modelling of Glisovic *et al.* (2012), the lateral stability and activity of this 'Siberian plume' is predicted to be long-lived, extending over several hundred million years. In general the amplitudes of the predicted upwellings are greater, and also more spatially focussed, than the more diffuse downwellings. These results are closely matched by the flow predictions obtained using the density anomalies in model $GyPSuM$ (Fig. 31).

The robustness and dynamical significance of the predicted lower-mantle superplumes may first be appreciated by noting that this long-wavelength pattern of upwellings is produced by an integral of the density anomalies across the entire depth of the mantle (see the kernels in Figs. 27, 28). This pattern of lower-mantle flow is therefore very well constrained by the joint seismic-geodynamic tomography inversions, particularly because they also include additional effects from lower-mantle compositional heterogeneity that acts to locally oppose or cancel thermally generated buoyancy. The dynamical significance of these lower-mantle superplumes, in terms of the long-term thermal evolution of the mantle, has been explored in detail by Glisovic *et al.* (2012). They carried out mantle convection simulations that were initiated with the present-day density heterogeneity in joint tomography model $TX2008$ and the thermal evolution of the 3-D structure in the mantle was integrated forwards over many hundred million years (in excess of 2 billion years, in total). These very-long time convection simulations revealed the extraordinary stability and longevity of the present-day superplumes shown in Fig. 31, requiring a strong flux ($> 13\,TW$) of heat entering the mantle from the core.

Arguably the most striking feature of the flow predictions shown in the upper and lower mantle (Figs. 30,31) is the strong, vertically coherent superplume-related upwelling under the EPR. This active upwelling, driven by the integrated buoyancy across the entire mantle, is clearly evident at all depths in the mantle. The EPR is thus the only oceanic ridge which sits above an upwelling that extends down to the CMB. This result that has not been found in previous tomography-based simulations of mantle dynamics because, until now, they employed purely seismic 3-D mantle models that have been converted *a-posteriori* to density with simple scaling factors. The longevity and lateral stability of the EPR superplume, possibly over several hundred million years as shown in Glisovic *et al.* (2012), should have a major impact on the long-term evolution of plate tectonics in the Pacific Ocean and for the overall thermal evolution of the mantle.

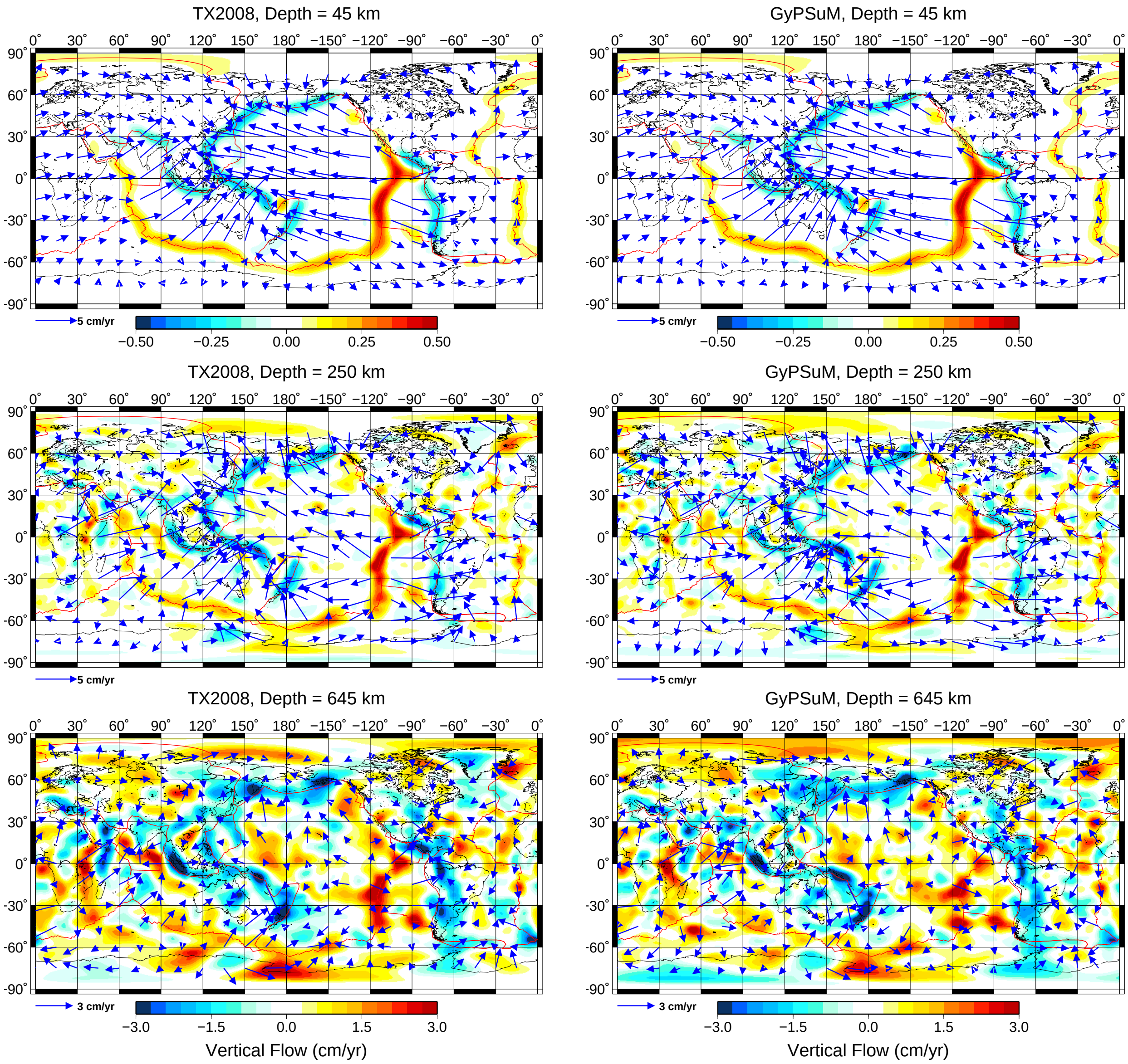
TX2008, Depth = 45 km
GyPSuM, Depth = 45 km
TX2008, Depth = 250 km
GyPSuM, Depth = 250 km
TX2008, Depth = 645 km
GyPSuM, Depth = 645 km
0° 30° 60° 90° 120° 150° 180° −150° −120° −90° −60° −30° 0°
90° 60° 30° 0° -30° -60° -90°
5 cm/yr
3 cm/yr
−0.50 −0.25 0.00 0.25 0.50
−3.0 −1.5 0.0 1.5 3.0
Vertical Flow (cm/yr)

Figure 30: Predicted mantle flow at three different depths in the upper mantle : lithosphere (top row), asthenosphere (middle row) and base of the transition zone (bottom row). The buoyancy forces are derived from two joint seismic-geodynamic inversions for mantle density anomalies: models $TX2008$ (Simmons *et al.* 2009) and $GyPSuM$ (Simmons *et al.* 2010) in the *left* and *right* columns, respectively. The viscous flow kernels are calculated on the basis of the 'V2' viscosity profile, derived from the joint GIA-convection inversions, shown in Fig. 9. The flow fields are all represented in terms of a spherical harmonic expansion truncated at degree $\ell = 72$. All flow predictions are in a mantle reference frame with zero global (net) rotation.

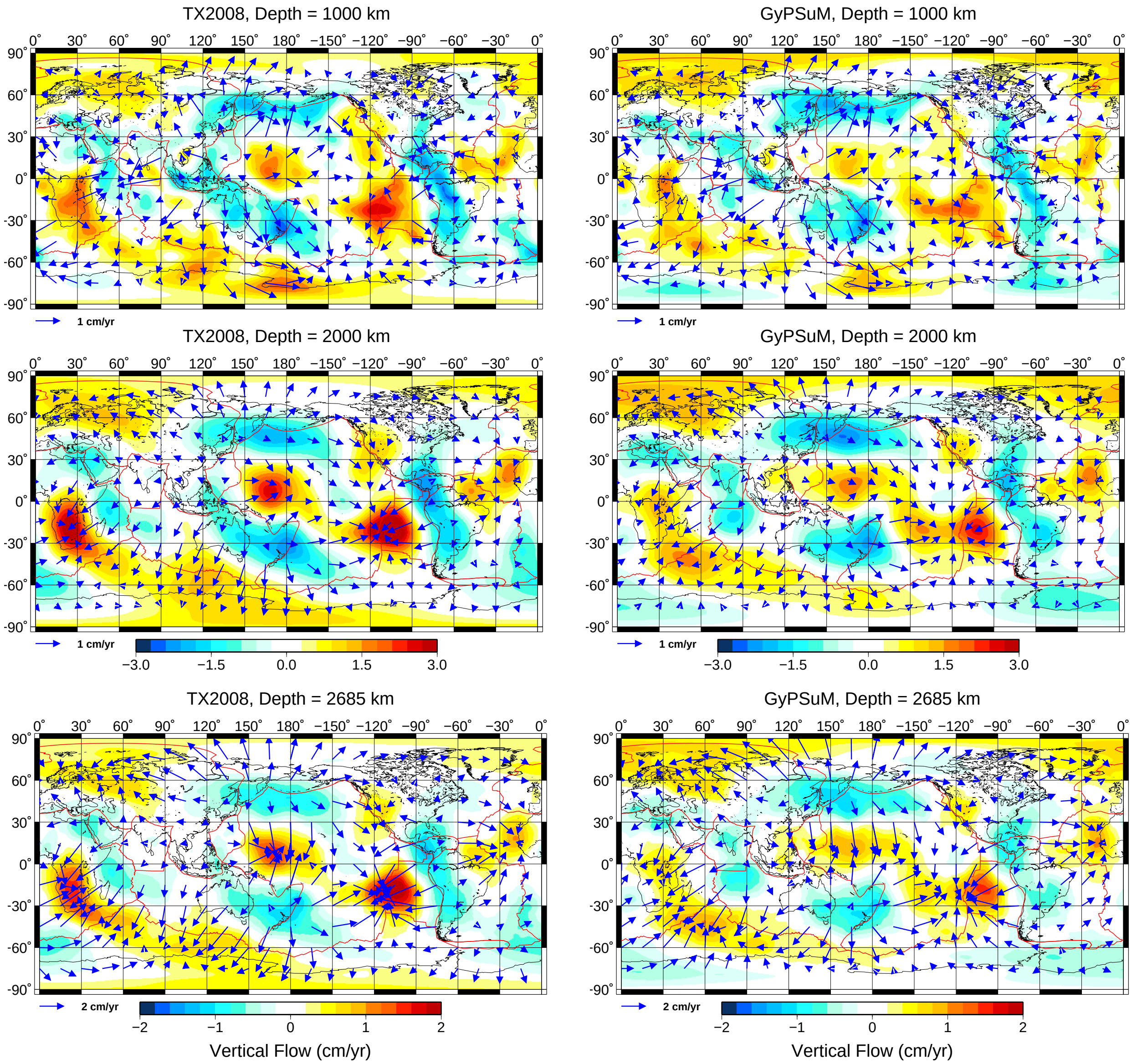
TX2008, Depth = 1000 km
GyPSuM, Depth = 1000 km
TX2008, Depth = 2000 km
GyPSuM, Depth = 2000 km
TX2008, Depth = 2685 km
GyPSuM, Depth = 2685 km
0° 30° 60° 90° 120° 150° 180° −150° −120° −90° −60° −30° 0°
90° 60° 30° 0° -30° -60° -90°
1 cm/yr
2 cm/yr
−3.0 −1.5 0.0 1.5 3.0
−2 −1 0 1 2
Vertical Flow (cm/yr)

Figure 31: Predicted mantle flow at three different depths in the lower mantle: 1000 km (top row), 2000 km (middle row) and 2685 km, corresponding to the top of the seismic D″-layer (bottom row). The flow calculation for input density anomalies from models $TX2008$ (Simmons *et al.* 2009) and $GyPSuM$ (Simmons *et al.* 2010) is identical to that in Fig. 30.

5 Concluding remarks

Tomography models of 3-D mantle structure, derived solely by inverting global seismic data, provide reasonably good fits to a wide variety of surface geodynamic observables related to mantle convection. The previous tomography-based models of mantle dynamics thus demonstrated that global seismic data can indeed be used to recover images of lateral heterogeneity that is produced by the thermal convection process in Earth's mantle.

The tabulated summaries of the predicted convection results (Tables 1 - 4) show however that the fits leave considerable room for improvement and this raises the fundamental question as to what aspects of the previous tomography models, or indeed the mantle flow calculations, may be deficient or inadequate.

A major development that goes far in addressing this outstanding question, is the technique for the joint or simultaneous inversion of global seismic and geodynamic data sets that also incorporates mineral physical constraints on the thermal contributions to mantle heterogeneity (e.g., Forte *et al.* 1994, Forte & Woodward 1997, Simmons et al.*et al.* 2006, 2007, 2009, 2010). These joint seismic-geodynamic inversions show that it is possible to obtain greatly improved fits to the global convection-related data (see Table 5) using the mantle flow theory which was presented and discussed in detail in section 2.3. The improved fits are found to be the result of relatively small adjustments (mainly in amplitude) to the seismically inferred heterogeneity in the mantle, especially the mid-mantle region, in order to satisfy the independent constraints on 3-D mantle structure imposed by the geodynamic data. Moreover, it is found that this reconciliation of seismic and geodynamic constraints can be achieved in the context of a dominantly thermal origin for mantle heterogeneity. The joint seismic-geodynamic inversions thus provide very strong support for the fundamental conclusion that most of the heterogeneity in the mantle is maintained by the process of thermal convection.

The greatly improved resolution of the 3-D density structure of the mantle provided by the joint inversions is yielding important new insights on the connections between the convective flow in the mantle and the surface geological evolution of our planet (e.g. Forte *et al.* 2010). Recent simulations of time-reversed convection using joint seismic-geodynamic tomography models also provide new perspectives on the evolution of 3-D mantle structure in the geological past and how this evolution controls changes in the surface topography of the Earth, with all the associated implications for global and regional changes in continental elevations and sea level (e.g. Moucha *et al.* 2008, Moucha & Forte 2011, Rowley *et al.* 2013). The key ingredient that ensures the reliability and realism of these time-reversed convection simulations is the use of density anomalies, coupled with appropriately constrained mantle viscosity (e.g. Fig. 9), derived from the joint seismic-geodyamic inversions. The use of geodynamically consistent tomography models ensures that the present-day mantle flow (see Figs. 30,31) is as realistic as possible. This is a fundamental issue, because the present-day mantle flow determines the critically important starting trajectory for the highly nonlinear time-reversed convection simulations.

Acknowledgements

I am indebted to a number of colleagues who have closely worked with me on many of the topics covered by this review. Their invaluable contributions are gratefully acknowledged: Petar Glisovic, Stephen Grand, Jerry Mitrovica, Robert Moucha, Sandrine Quéré, David Rowley and Nathan Simmons. This work has been made possible thanks to support provided by the Canadian Institute for Advanced Research and the Canadian National Science and Engineering Research Council.

References

[1] Alisic, L., Gurnis, M., Stadler, G., Burstedde, C. & Ghattas, O., Multi-scale dynamics and rheology of mantle flow with plates, *J. Geophys. Res.* , **117**, pp 1–27, B10402, doi:10.1029/2012JB009234, 2012.

[2] Antolik, M., Gu, Yu J., Ekström, G. & Dziewonski, A.M., J362D28; a new joint model of compressional and shear velocity in the Earth's mantle, *Geophys. J. Int.* , **153**, 443–466, 2003.

[3] Bassin, C., Laske, G. & Masters, G., The current limits of resolution for surface wave tomography in North America, *Eos Trans. AGU* , **81**(48), 897, 2000.

[4] Becker, T.W. & Faccenna, C., Mantle conveyor beneath the Tethyan collisional belt, *Earth Planet. Sci. Lett.* , **310**, 453–461, 2011.

[5] Becker, T.W. & Boschi, L., A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.*, **3**, 10.129/2001GC000168, 2002.

[6] Boschi, L. & Dziewonski, A.M., Whole Earth tomography from delay times of P, PcP, and PKP phases: Lateral heterogeneities in the outer core or radial anisotropy in the mantle?, *J. Geophys. Res.* , **105**, 13,675-13,696, 2000.

[7] Bunge, H.-P., Richards, M.A. & Baumgardner, J.R., Effect of depth-dependent viscosity on the planform of mantle convection, *Nature*, **379**, 436–438, 1996.

[8] Bunge, H.-P., M.A. Richards, C. Lithgow-Bertelloni, J.R. Baumgardner, S.P. Grand & B.A. Romanowicz, Time scales and heterogeneous structure in geodynamic Earth models, *Science* , **280**, 91–95, 1998.

[9] Čadek, O. & Fleitout, L., Effect of lateral viscosity variations in the top 300 km on the geoid and dynamic topography, *Geophys. J. Int.* , **152**, 566–580, 2003.

[10] Čadek, O., Yuen, D. A., Steinbach, V., Chopelas, A. & Matyska, C., Lower mantle thermal structure deduced from seismic tomography, mineral physics and numerical modelling. *Earth Planet. Sci. Lett.* , **121**, 385–402, 1994.

[11] Cammarano, F., Goes, S., Vacher, P. & Giardini, D., Inferring upper-mantle temperatures from seismic velocities, *Phys. Earth Planet. Inter.* , **138**, 197–222, 2003.

[12] Cathles, L.M., *The Viscosity of the Earth's Mantle*, Princeton Univ. Press, Princeton, N.J., 1975.

[13] Chase, C.G., Subduction, the geoid, and lower mantle convection, *Nature* , **282**, 464–468, 1979.

[14] Chase, C.G. & Sprowl, D.R., The modern geoid and ancient plate boundaries, *Earth Planet. Sci. Lett.* , **62**, 314–320, 1983.

[15] Clayton, R.W. & R.P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times, *Eos Trans. AGU* , **64**(45), 776, 1983.

[16] Conrad, C.P., Bianco, T.A., Smith, E.I. & Wessel, P., Patterns of intraplate volcanism controlled by asthenospheric shear, *Nature Geoscience*, **4**, 317–321, 2011.

[17] Constable, S. C., Parker, R. L. & Constable, C. G., Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**, 289–300, 1987

[18] Corrieu, V., Ricard, Y. & Froidevaux, C., Converting mantle tomography into mass anomalies to predict the Earth's radial viscosity. *Phys. Earth Planet. Inter.* , **84**, 3–13, 1994.

[19] Corrieu, V., C. Thoraval, & Y. Ricard, Mantle dynamics and geoid Green functions, *Geophys. J. Int.* , **120**, 516–523, 1995.

[20] Crosby, A.G., McKenzie, D. & Sclater, J.G., The relationship between depth, age and gravity in the oceans, *Geophys. J. Int.* , **166**, 553–573, 2006.

[21] Davies, D.R., Goes, S., Davies, J.H., Schuberth, B.S.A., Bunge, H.-P. & Ritsema, J., Reconciling dynamic and seismic models of Earth?s lower mantle: The dominant role of thermal heterogeneity, *Earth Planet. Sci. Lett.* , **353-354**, 253–269, 2012.

[22] DeMets, C. R., Gordon, R. G., Argus, D. F. & Stein, S., Current plate motions. *Geophys. J. Int.* **101**, 425–478, 1990.

[23] Defraigne, P., Internal loading of a compressible Earth: Effect of a solid lithosphere *Phys. Earth Planet. Inter.* , **101**, 303–313, 1997.

[24] Deschamps, F., Snieder, R. & Trampert, J., The relative density-to-shear velocity scaling in the uppermost mantle. *Phys. Earth Planet. Inter.* , **124**, 193–212, 2001

[25] deWit, R.W.L., Trampert, J. & van der Hilst, R.D., Toward quantifying uncertainty in travel time tomography using the null-space shuttle, *J. Geophys. Res.* , **117**, B03301, doi:10.1029/2011JB008754, 2012.

[26] Doin, M-P & Fleitout, L., Flattening of the oceanic topography and geoid: thermal versus dynamic origin, *Geophys. J. Int.* , **143**, 582–594, 2000.

[27] Dziewonski, A. M., Mapping the lower mantle: determination of lateral heterogeneity in P velocity up to degree and order 6, *J. Geophys. Res.* , **89**, 5929–5952, 1984.

[28] Dziewonski, A.M. & Anderson, D.L., Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356, 1981.

[29] Dziewonski, A.M. & Romanowicz, B.A., Overview, in Volume 1 of *Treatise of Geophysics*, edited by B.A. Romanowicz & A.M. Dziewonski, Elsevier, pp. 1–29, 2007.

[30] Dziewonski, A.M., B.H. Hager & R.J. O'Connell, Large scale heterogeneity in the lower mantle, *J. Geophys. Res.* , **82**, 239–255, 1977.

[31] Ekström, G. & Dziewonski, A.M., The unique anisotropy of the Pacific upper mantle. *Nature* , **394**, 168–172, 1998.

[32] Forte, A.M., Seismic-geodynamic constraints on mantle flow: Implications for layered convection, mantle viscosity, and seismic anisotropy in the deep mantle. In *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale* (ed. S.–i. Karato *et al.*). Geophys. Monogr. Ser., **117**, pp. 3–36, Washington DC: American Geophysical Union, 2000.

[33] Forte, A.M., Constraints on seismic models from other disciplines – Implications for mantle dynamics and composition, in Volume 1 of *Treatise of Geophysics*, edited by B. Romanowicz & A.M. Dziewonski, Elsevier, pp. 805–858, 2007.

[34] Forte, A.M. & Peltier, W.R., Plate tectonics and aspherical Earth structure: The importance of poloidal–toroidal coupling. *J. Geophys. Res.* , **92**, 3645–3679, 1987.

[35] Forte, A.M. & Peltier, W.R., Core-mantle boundary topography and whole-mantle convection, *Geophys. Res. Lett.* , **16**, 621–624, 1989.

[36] Forte, A.M. & Peltier, W.R., Viscous flow models of global geophysical observables,1 Forward problems. *J. Geophys. Res.* **96**, 20131–20159, 1991.

[37] Forte, A.M. & Peltier, W.R., The kinematics and dynamics of poloidal-toroidal coupling in mantle flow: The importance of surface plates and lateral viscosity variations. *Adv. Geophys.* **36**, 1–119, 1994.

[38] Forte, A.M. & Mitrovica, J.X., New inferences of mantle viscosity from joint inversion of long-wavelength mantle convection and post–glacial rebound data, *Geophys. Res. Lett.* , **23**, 1147–1150, 1996.

[39] Forte, A.M. & R.L. Woodward, Seismic-geodynamic constraints on vertical flow between the upper and lower mantle: The dynamics of the 670 km seismic discontinuity, in *The Fluid Dynamics of Astrophysics and Geophysics*, vol. 7, *Earth's Deep Interior*, edited by D. Crossley, pp. 337–404, Gordon and Breach, Newark, N.J., 1997a.

[40] Forte, A.M. & R.L. Woodward, Seismic-geodynamic constraints on three-dimensional structure, vertical flow, and heat transfer in the mantle, *J. Geophys. Res.* , **102**, 17981–17994, 1997b.

[41] Forte, A.M. & Perry, H.K.C., Geodynamic evidence for a chemically depleted continental tectosphere, *Science* , **290**, 1940–1944, 2000.

[42] Forte, A.M., & J.X. Mitrovica, Deep-mantle high-viscosity flow and thermochemical structure inferred from seismic and geodynamic data, *Nature* , **410**, 1049–1056, 2001.

[43] Forte, A.M., Peltier, W.R. & Dziewonski, A.M., Inferences of mantle viscosity from tectonic plate velocities. *Geophys. Res. Lett.* **18**, 1747–1750, 1991.

[44] Forte, A.M., Dziewonski, A.M. & Woodward, R.L., Aspherical structure of the mantle, tectonic plate motions, nonhydrostatic geoid, and topography of the core-mantle boundary. In *Dynamics of the Earth's Deep Interior and Earth Rotation* (eds. J.–L. LeMouël, D. E. Smylie & T. Herring). Geophys. Monogr. Ser., **72**, pp. 135–166, Washington DC: American Geophysical Union, 1993a.

[45] Forte, A.M., W.R. Peltier, A.M. Dziewonski & R.L. Woodward, Dynamic surface topography: A new interpretation based upon mantle flow models derived from seismic tomography, *Geophys. Res. Lett.* , **20**, 225–228, 1993b.

[46] Forte, A.M., W.R. Peltier, A.M. Dziewonski & R.L. Woodward, Reply to comment by M. Gurnis,*Geophys. Res. Lett.* **20**, 1665–1666, 1993c.

[47] Forte, A.M., Woodward, R.L. & Dziewonski, A.M., Joint inversions of seismic and geodynamic data for models of three-dimensional mantle heterogeneity. *J. Geophys. Res.* **99**, 21857–21877, 1994

[48] Forte, A.M., Mitrovica, J.X. & Woodward, R.L., Seismic-geodynamic determination of the origin of excess ellipticity of the core-mantle boundary, *Geophys. Res. Lett.* **22**, 1013–1016, 1995a.

[49] Forte, A. M., A. M. Dziewonski & R. J. O'Connell, Continent-ocean chemical heterogeneity in the mantle based on seismic tomography, *Science* , **268**, 386–388, 1995b.

[50] Forte, A.M., J.X. Mitrovica, & A. Espesset, Geodynamic and seismic constraints on the thermochemical structure and dynamics of convection in the deep mantle, *Philos. Trans. R. Soc. London A* , **360**, 2521–2543, 2002.

[51] Forte, A.M., R. Moucha, D.B. Rowley, S. Quéré, J.X. Mitrovica, N.A. Simmons & S.P. Grand, Recent tectonic plate decelerations driven by mantle convection, *Geophys. Res. Lett.* , **36**, L23301, doi:10.1029/2009GL040224, 2009.

[52] Forte, A.M., S. Quéré, R. Moucha, N.A. Simmons, S.P. Grand, J.X. Mitrovica & Rowley, D.B., Joint seismic-geodynamic-mineral physical modelling of African geodynamics: A reconciliation of deep-mantle convection with surface geophysical constraints, *Earth Planet. Sci. Lett.* , **295**, 329--341, 2010.

[53] Gable, C.W., O'Connell, R.J. & Travis, B.J., Convection in three dimensions with surface plates. *J. Geophys. Res.* **96**, 8391–8405, 1991.

[54] Garcia, R.& Souriau, A., Amplitude of the core-mantle boundary topography estimated by stochastic analysis of core phases, *Phys. Earth Planet. Inter.* , **117**, 345–359, 2000.

[55] Glisovic, P., Forte, A.M. & Moucha, R., Time-dependent convection models of mantle thermal structure constrained by seismic tomography and geodynamics: implications for mantle plume dynamics and CMB heat flux, *Geophys. J. Int.* , **190**, 785–815, 2012.

[56] Gordon, R.G. & Jurdy, D.M., Cenozoic global plate motions, *J. Geophys. Res.* , **91**, 12389–12406, 1986.

[57] Grand, S.P., Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.* , **99**, 11591–11621, 1994.

[58] Grand, S.P., Mantle shear-wave tomography and the fate of subducted slabs, *Philos. Trans. R. Soc. London A* , **360**, 2475–2491, 2002.

[59] Gurnis, M., Bounds on global dynamic topography from Phanerozoic flooding of continental platforms, *Nature* , **344**, 754–756, 1990.

[60] Gurnis, M., Comment, *Geophys. Res. Lett.* , **20**, 1663–1664, 1993.

[61] Gwinn, C.R., Herring, T.A., & Shapiro, I.I., Geodesy by radio interferometry; studies of the forced nutations of the Earth; 2, Interpretation, *J. Geophys. Res.* , **91**, 4755–4766, 1986.

[62] Hager, B.H., Subducted slabs and the geoid: Constraints on mantle rheology and flow, *J. Geophys. Res.* , **89**, 6003–6015, 1984.

[63] Hager, B.H. & O'Connell, R.J., A simple global model of plate dynamics and mantle convection. *J. Geophys. Res.* **86**, 4843–4867, 1981.

[64] Hager, B.H., & M.A. Richards, Long-wavelength variations in the Earth's geoid: physical models and dynamical implications, *Philos. Trans. R. Soc. London A* , **328**, 309–327, 1989.

[65] Hager, B. H. & Clayton, R. W., Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid. In *Mantle Convection: Plate tectonics and global dynamics* (ed. W. R. Peltier), pp. 657–763. New York: Gordon and Breach, 1989.

[66] Hager, B.H., R.W. Clayton, M.A. Richards, R.P. Comer & A.M. Dziewonski, Lower mantle heterogeneity, dynamic topography and the geoid, *Nature* , **313**, 541–545, 1985.

[67] Haskell, N.A., The motion of a viscous fluid under a surface load, *Physics* (N.Y.), **6**, 265–269, 1935.

[68] Herring, T.A., Gwinn, C.R. & Shapiro, I.I., Geodesy by radio interferometry; studies of the forced nutations of the Earth; 1, Data analysis, *J. Geophys. Res.* , **91**, 4745–4754, 1986.

[69] Jackson, J.D., *Classical Electrodynamics*, J. Wyley & Sons, Inc., NY, USA, 1975.

[70] Jackson, I., Chemical composition and temperature of the lower mantle from seismological models: Residual uncertainties. In *Integrated models of Earth structure and evolution, AGU Virtual Meeting - 2001 Spring Meeting, 20 June,* [online], 2001.

[71] Jarvis, G.T., & Peltier, W.R., Mantle convection as a boundary layer phenomenon, *Geophys. J. R. Astr. Soc.* **68**, 385–424, 1982.

[72] Jeanloz, R. & Thompson, A.B., Phase transitions and mantle discontinuities, *Rev. Geophys. Space Phys.* , **21**, 51–74, 1983.

[73] Jordan, T.H., Continents as a chemical boundary layer, *Philos. Trans. R. Soc. London A* , **301**, 359–373, 1981.

[74] Karato, S.–i., Importance of anelasticity in the interpretation of seismic tomography. *Geophys. Res. Lett.* , **20**, 1623–1626, 1993.

[75] Karato, S.-I. & Karki, B.B., Origin of lateral variation of seismic wave velocities and density in the deep mantle, *J. Geophys. Res.* , **106**, 21771–21784, 2001.

[76] Karpychev, M. & Fleitout, L., Long-wavelength geoid: the effect of continental roots and lithosphere thickness variations. *Geophys. J. Int.* , **143**, 945–963, 2000.

[77] Kaufmann, G. & Lambeck, K., Mantle dynamics, postglacial rebound and the radial viscosity profile, *Phys. Earth Planet. Inter.* , **121**, 301-324, 2000.

[78] Kido, M., Yuen, D.A., Čadek, O. & Nakakuki, T., Mantle viscosity derived by genetic algorithm using oceanic geoid and seismic tomography for whole-mantle versus blocked-flow situations, *Phys. Earth Planet. Inter.* , **107**, 307–326, 1998.

[79] King, S., & G.Masters, An inversion for the radial viscosity structure using seismic tomography, *Geophys. Res. Lett.* , **19**, 1551–1554, 1992.

[80] King, S.D., Models of mantle viscosity, in *A Handbook of Physical Constants: Mineral Physics and Crystallography*, edited by T.J. Ahrens, pp. 227–236 AGU Reference Shelf Series, vol. 2, American Geophysical Union (Washington DC), 1995.

[81] Kustowski, B., Ekström, G. & Dziewonski, A.M., Anisotropic shear-wave velocity structure of the Earth's mantle: A global model, *J. Geophys. Res.* , **113**, B06306, doi:10.1029/2007JB005169, 2008.

[82] Lambeck, K., Johnston, P., Smither, C. & Nakada, M., Glacial rebound of the British Isles. III. Constraints on mantle viscosity. *Geophys. J. Int.* , **125**, 340-354, 1996.

[83] Lambeck, K., Smither, C. & Johnston, P., Sea-level change, glacial rebound and mantle viscosity for northern Europe. *Geophys. J. Int.* , **134**, 102-144, 1998.

[84] Landau, L.D., and E.M. Lifshitz, *Fluid Mechanics*, vol.**6**, Pergamon (New York), 1959.

[85] Lemoine, F., N. Pavlis, S. Kenyon, R. Rapp, E. Pavlis, & B. Chao (1998), New high-resolution model developed for earth's gravitational field, *Eos Trans. AGU* , **79**(9), 113, 1998.

[86] Le Stunff, Y., & Y. Ricard, Topgraphy and geoid due to lithospheric mass anomalies, *Geophys. J. Int.* , **122**, 982–990, 1995.

[87] Le Stunff, Y., & Y. Ricard, Partial advection of equidensity surfaces; a solution for the dynamic topography problem?, *J. Geophys. Res.* , **102**, 24,655–24,667, 1997.

[88] Levander, A., Schmandt, B., Miller, M.S., Liu, K., Karlstrom, K.E., Crow, R.S., Lee, C.-T.A. & E. D. Humphreys, E.D., Continuing Colorado plateau uplift by delamination-style convective lithospheric downwelling, *Nature* , **472**, 461–465, 2011.

[89] Li, C., van der Hilst, R.D., Engdahl, E.R. & Burdick, S., A new global model for P wave speed variations in Earth's mantle, *Geochem. Geophys. Geosyst.*, **9**, Q05018, doi:10.1029/2007GC001806, 2008.

[90] Lithgow-Bertelloni, C. & Gurnis, M., Cenozoic subsidence and uplift of continents from time-varying dynamic topography *Geology*, **25**, 735–738, 1997.

[91] Lithgow-Bertelloni, C. & Richards, M.A., The dynamics of Cenozoic and Mesozoic plate motions. *Rev. Geophys.* **36**, 27–78, 1998.

[92] Martinec, Z., Matyska, C., Cadek O. & Hrdina P., The Stokes problem with 3-D Newtonian rheology in a sphericall shell. *Computer Phys. Commun.*, **76**, 63–79, 1993.

[93] Masters, G., Laske, G., Bolton, H. & Dziewonski, A.M., The relative behaviour of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure. In *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale* (ed. S.–i. Karato *et al.*). Geophys. Monogr. Ser., **117**, pp. 63–87, Washington DC: American Geophysical Union, 2000.

[94] Mathews, P.M., T. A. Herring & B. A. Buffett, Modeling of nutation and precession: New nutation series for nonrigid Earth and insights into the Earth's interior, *J. Geophys. Res.* , **107**, 2068, doi:10.1029/2001JB000390, 2002.

[95] McNamara, A.K. & S.-J. Zhong, Thermochemical structures beneath Africa and the Pacific Ocean, *Nature*, **437**, 1136–1139, 2005.

[96] McConnell, R.K., Viscosity of the mantle from relaxation time spectra of isostatic adjustment, *J. Geophys. Res.* , **73**, 7089–7105, 1968.

[97] Mégnin, C. & Romanowicz, B., The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms. *Geophys. J. Int.* , **143**, 709–728, 2000.

[98] Mitrovica, J.X., Haskell[1935] revisited, *J. Geophys. Res.* , **101**, 555-569, 1996.

[99] Mitrovica, J.X. & Forte, A.M., Radial profile of mantle viscosity: results from the joint inversion of convection and post-glacial rebound observables. *J. Geophys. Res.* , **102**, 2751–2769, 1997.

[100] Mitrovica, J.X. & Forte, A.M., A new inference of mantle viscosity based upon joint inversion of convection and glacial isostatic adjustment data, *Earth Planet. Sci. Lett.* , **225**, 177–189, 2004.

[101] Montelli, R., Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R. & Hung, S.H., Finite-frequency tomography reveals a variety of plumes in the mantle, *Science* , **303**, 338–343, 2004.

[102] Morelli, A., & A.M. Dziewonski, Topography of the core-mantle boundary and lateral homogeneity of the liquid core, *Nature* , **325**, 678–683, 1987.

[103] Mosca, I., Cobden, L., Deuss, A., Ritsema, J. & Trampert, J., Seismic and mineralogical structures of the lower mantle from probabilistic tomography, *J. Geophys. Res.* , **117**, B06304, doi:10.1029/2011JB008851, 2012.

[104] Moucha, R., Forte, A.M., Mitrovica, J.X. & Daradich A., Geodynamic implications of lateral variations in mantle rheology on convection related observables and inferred viscosity models, *Geophys. J. Int.* , **169**, 113–135, 2007.

[105] Moucha, R., A.M. Forte, J.X. Mitrovica, D.B. Rowley, S. Quéré, N.A. Simmons & S.P. Grand, Dynamic topography and long-term sea-level variations: There is no such thing as a stable continental platform, *Earth Planet. Sci. Lett.* , **271**, 101-108, 2008.

[106] Moucha, R. & Forte, A.M., Changes in African topography driven by mantle convection, *Nature Geoscience*, **4**, 707–712, 2011.

[107] Muller, R.D., Roest, W.R., Royer, J.-Y., Gahagan, L.M. & Sclater, J.G., Digital isochrons of the world's ocean floor, *J. Geophys. Res.* , **102**, 3211–3214, 1997.

[108] Nakada, M.& Lambeck, K., Late Pleistocene and Holocene sea-level change in the Australian region and mantle rheology, *Geophys. J. Int.*, **96**, 497–517, 1989.

[109] Nakagawa, T. & Tackley, P.J., Lateral variations in CMB heat flux and deep mantle seismic velocity caused by a thermal?chemical-phase boundary layer in 3D spherical convection, *Earth Planet. Sci. Lett.* , **271**, 348–358, 2008.

[110] Nakagawa, T., Tackley,P.J., Deschamps, F. & Connolly, J.A.D., The influence of MORB and Harzburgite composition on thermo-chemical mantle convection in a 3D spherical shell with self-consistently calculated mineral physics, *Earth Planet. Sci. Lett.* , **296**, 403–412, 2010.

[111] Nicolas, A., & Poirier, J.-P., *Crystalline Plasticity and Solid State Flow in Metamorphic Rocks*, John Wiley and Sons, London, UK, pp 444, 1976.

[112] Nolet, G., Karato, S.–I. & Montelli, R., Plume fluxes from seismic tomography, *Earth Planet. Sci. Lett.* , **248**, 685–699, 2006.

[113] O'Connell, R. J., Pleistocene glaciation and viscosity of the lower mantle, *Geophys. J. R. Astron. Soc.* , **23**, 299–327, 1971.

[114] Obayashi, M. & Fukao, Y., P and PcP travel time tomography for the core-mantle boundary, *J. Geophys. Res.* , **102**, 17,825–17,841, 1997.

[115] Oganov, A. R., Brodholt, J. P. & Price, G. D., *Ab initio* elasticity and thermal equation of state of $MgSiO_3$ perovskite. *Earth Planet. Sci. Lett.* , **184**, 555–560, 2001a.

[116] Oganov, A. R., Brodholt, J. P. & Price, G. D., The elastic constants of $MgSiO_3$ perovskite at pressures and temperatures of the Earth's mantle. *Nature* , **411**, 934–937, 2001b

[117] Paige, C.C. & Saunders, M.A., LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software*, **8**, 43–71, 1982.

[118] Panasyuk, S.V. & Hager, B.H., Inversion for mantle viscosity profiles constrained by dynamic topography and the geoid, and their estimated errors, *Geophys. J. Int.* , **143**, 821–836, 2000.

[119] Panasyuk, S.V., Hager, B.H. & Forte, A.M., Understanding the effects of mantle compressibility on geoid kernels, *Geophys. J. Int.* , **124**, 121–133, 1996.

[120] Pari, G., Crust 5.1-based inference of the Earth's dynamic surface topography: geodynamic implications. *Geophys. J. Int.* **144**, 501–516, 2001.

[121] Pari, G. & Peltier, W.R., Subcontinental mantle dynamics; a further analysis based on the joint constraints of dynamic surface topography and free-air gravity, *J. Geophys. Res.* , **105**, 5635–5662, 2000.

[122] Parsons, B. & Sclater, J.G. An analysis of the variation of ocean floor bathymetry and heat flow with age, *J. Geophys. Res.* , **82**, 803–827, 1977.

[123] Parsons, B. & McKenzie, D. Mantle convection and the thermal structure of the plates, *J. Geophys. Res.* , **83**, 4485–4496, 1978.

[124] Parsons, B. & Daly, S., The relationship between surface topography, gravity anomalies, and temperature structure of convection, *J. Geophys. Res.* , **88**, 1129–1144, 1983.

[125] Pekeris, C.L., Thermal convection in the interior of the Earth, *Mon. Not. R. Astron. Soc. Geophys. Suppl*, **3**, 343–367, 1935.

[126] Peltier, W.R., New constraint on transient lower mantle rheology and internal mantle buoyancy from glacial rebound data, *Nature*, **318**, 614–617, 1985.

[127] Peltier, W.R. & J.T. Andrews, Glacial isostatic adjustment, I, The forward problem, *Geophys. J. R. Astron. Soc.*, **46**, 605–646, 1976.

[128] Perry, H.K.C., Forte, A.M. & D.W.S. Eaton, Upper-mantle thermochemical structure below North America from seismicgeodynamic flow models, *Geophys. J. Int.* , **154**, 279–299, 2003.

[129] Phinney, R.A. & R. Burridge, Representation of the elastic-gravitational excitation of a spherical Earth model by generalized spherical harmonics, *Geophys. J. R. Astron. Soc.* , **34**, 451–487, 1973.

[130] Quéré, S. & A.M. Forte, Influence of past and present-day plate motions on spherical models of mantle convection: Implications for mantle plumes and hotspots, *Geophys. J. Int.*, **165**, 1041–1057, 2006.

[131] Ranalli, G. Mantle rheology: Radial and lateral viscosity variations inferred from microphysical creep laws, *J. Geodynamics*, **32**, 425-444, 2001.

[132] Ricard, Y. & Vigny, C., Mantle dynamics with induced plate tectonics, *J. Geophys. Res.* , **94**, 17543–17559, 1989.

[133] Ricard, Y. & Wuming, B., Inferring viscosity and the 3-D density structure of the mantle from geoid, topography and plate velocities, *Geophys. J. Int.* , **105**, 561–572, 1991.

[134] Ricard, Y., Fleitout, L. & Froidevaux, C., Geoid heights and lithospheric stresses for a dynamic Earth, *Ann. Geophys.*, **2**, 267–286, 1984.

[135] Ricard, Y., Froidevaux, C. & Fleitout, L., Global plate motion and the geoid: A physical model, *Geophys. J. Int.* , **93**, 477–484, 1988.

[136] Ricard, Y., Vigny, C. & Froidevaux, C., Mantle heterogeneities, geoid, and plane motion – A Monte Carlo inversion, *J. Geophys. Res.* , **94**, 13739–13754, 1989.

[137] Ricard, Y., Richards, M., Lithgow-Bertelloni, C.& Le Stunff, Y., A geodynamic model of mantle density heterogeneity, *J. Geophys. Res.* , **98**, 21895–21909, 1993.

[138] Ricard, Y., Chambat F. & Lithgow-Bertelloni, C., Gravity observations and 3D structure of the Earth, *Comptes Rendus Geoscience*, in press, 2006.

[139] Richards, M.A., Hot spots and the case for a high viscosity lower mantle, in *Glacial Isostasy, Sea-Level and Mantle Rheology*, *NATO ASI Ser. C*, vol. 334, edited by R. Sabadini, K. Lambeck & E. Boschi, pp. 571–588, Kluwer Acad., 1991.

[140] Richards, M.A., & Hager, B.H., Geoid anomalies in a dynamic Earth. *J. Geophys. Res.* , **89**, 5987–6002, 1984.

[141] Richards, M.A., & Hager, B.H., Effects of lateral viscosity variations on long-wavelength geoid anomalies and topography, *J. Geophys. Res.* , **94**, 10299–10313, 1989.

[142] Richards, M.A. & Engebretson, D.C., Large-scale mantle convection and the history of subduction, *Nature* , **355**, 437–440, 1992.

[143] Richards, M.A., Y. Ricard, C. Lithgow-Bertelloni, G. Spada & R. Sabadini, An explanation for the long-term stability of Earth's rotation axis, *Science*, **275**, 372–375, 1997.

[144] Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., Complex shear velocity structure imaged beneath Africa and Iceland. *Science* , **286**, 1925–1928, 1999.

[145] Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.* , **184**, 1223–1236, 2011.

[146] Ritzwoller, M.H., Shapiro, N.M. & Zhong, S.-J., Cooling history of the Pacific lithosphere, *Earth Planet. Sci. Lett.* , **226**, 69–84, 2004.

[147] Rodgers, A. & Wahr, J., Inference of core-mantle boundary topography from ISC PcP and PKP traveltimes, *Geophys. J. Int.* , **115**, 991–1011, 1993.

[148] Röhm, A.H.E., Snieder, R., Goes, S. & Trampert, J., Thermal structure of continental upper mantle inferred from S-wave velocity and surface heat flow, *Earth Planet. Sci. Lett.* , **181**, 395-407, 2000.

[149] Romanowicz, B., Global mantle tomography: Progress status in the past 10 years, *Annu. Rev. Earth Planet. Sci.* , **31**, 303–328, 2003.

[150] Rowley, D.B., Forte, A.M., Moucha, R., Mitrovica, J.X., Simmons, N.A. & Grand, S.P., Dynamic Topography Change of the Eastern U. S. Since 3 Million Years Ago, *Science*, **340**, 1560–1563, 2013.

[151] Sabadini, R., Yuen, D.A. & Gasperini, P. The effects of transient rheology on the interpretation of lower mantle rheology, *Geophys. Res. Lett.*, **12**, 361–365, 1985.

[152] Sabadini, R. & Yuen, D.A., Mantle stratification and long-term polar wander, *Nature*, **339**, 373–375, 1989.

[153] Sammis, C.G., Smith, J.C., Schubert, G. & Yuen, D.A., Viscosity-depth profile of the earth's mantle: Effects of polymorphic phase transitions, *J. Geophys. Res.* , **82**, 3747–3761, 1977.

[154] Schuberth, B.S.A., Bunge, H.–P. & Ritsema, J., Tomographic filtering of high-resolution mantle circulation models: Can seismic heterogeneity be explained by temperature alone?, *Geochem. Geophys. Geosyst.*, **10**, Q05W03, doi:10.1029/2009GC002401, 2009.

[155] Schuberth, B.S.A., Zaroli, C. & Nolet, G., Synthetic seismograms for a synthetic Earth: long-period P- and S-wave traveltime variations can be explained by temperature alone, *Geophys. J. Int.* , **188**, 1393–1412, 2012.

[156] Shen, W., Ritzwoller, M.H. & Schulte-Pelkum, V., A 3-D model of the crust and uppermost mantle beneath the Central and Western US by joint inversion of receiver functions and surface wave dispersion, *J. Geophys. Res.* , **118**, 262–276, doi:10.1029/2012JB009602, 2013.

[157] Simmons, N.A., Forte, A.M. & Grand, S.P., Constraining modes of mantle flow with seismic and geodynamic data: A joint approach, *Earth Planet. Sci. Lett.* , **246**, 109-124, 2006.

[158] Simmons, N.A., *Mantle Heterogeneity and Flow from Seismic and Geodynamic Constraints*, Doctoral dissertation, University of Texas at Austin, 2007.

[159] Simmons, N.A., Forte, A.M. & Grand, S.P., Thermochemical structure and dynamics of the African superplume, *Geophys. Res. Lett.* , **34**, L02301, doi:10.1029/2006GL028009, 2007.

[160] Simmons, N.A., Forte, A.M. & Grand, S.P., Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity: Implications for the relative importance of thermal versus compositional heterogeneity, *Geophys. J. Int.* , **177**(3), 1284–1304, 2009.

[161] Simmons, N.A., Forte, A.M., Boschi, L. & Grand, S.P., GyPSuM: A joint tomographic model of mantle density and seismic wave speeds, *J. Geophys. Res.* , **115**, doi:10.1029/2010JB007631, 2010.

[162] Smirnov, A.V. & Tarduno, J.A., Co-location of eruption sites of the Siberian Traps and North Atlantic Igneous Province: Implications for the nature of hotspots and mantle plumes, *Earth Planet. Sci. Lett.* , **297**(3-4), 687–690, 2010.

[163] Sobolev, S. V., Zeyen, H., Granet, M., Achauer, U., Bauer, C., Werling, F., Altherr, R. & Fuchs, K., Upper mantle temperatures and lithosphere-asthenosphere system beneath the French Massif Central constrained by seismic, gravity, petrologic and thermal observations. *Tectonophysics* **275**, 143–164, 1997.

[164] Spada, G., Ricard, Y. & Sabadini, R., Excitation of true polar wander by subduction, *Nature*, **360**, 452–454, 1992.

[165] Stacey, F.D., Thermoelasticity of a mineral composite and a reconsideration of lower mantle properties. *Phys. Earth Planet. Inter.* , **106**, 219–236, 1998.

[166] Stein, C.A. & Stein, S., A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature* , **359**, 123–129, 1992.

[167] Steinberger, B. & O'Connell, R.J., Change of the Earth's rotation axis inferred from the advection of mantle density heterogeneities, *Nature*, **387**, 169–173, 1997.

[168] Stixrude, L. & Lithgow-Bertelloni, C., The origin of lateral heterogeneity in the mantle. In *Integrated models of Earth structure and evolution, AGU Virtual Meeting - 2001 Spring Meeting, 20 June,* [online], 2001.

[169] Stixrude, L. & Lithgow-Bertelloni, C., Mineralogy and elasticity of the oceanic upper mantle: Origin of the low-velocity zone, *J. Geophys. Res.* , **110**, B03204, doi:10.1029/2004JB002965, 2005.

[170] Su, W.-J., R.L. Woodward, & A.M. Dziewonski, Deep origin of mid-ocean ridge seismic velocity anomalies, *Nature*, **360**, 149–152, 1992.

[171] Sze, E.K.M. & van der Hilst, R.D., Core mantle boundary topography from short period PcP, PKP, and PKKP data, *Phys. Earth Planet. Inter.* , **135**, 27–46, 2003.

[172] Thoraval, C. & Richards, M.A., The geoid constraint in global geodynamics: Viscosity structure, mantle heterogeneity models, and boundary conditions, *Geophys. J. Int.* , **131**, 1–8, 1997.

[173] Thoraval, C., Machetel, P. & Cazenave, A., Locally layered convection inferred from dynamic models of the Earth's mantle, *Nature* , **375**, 777–780, 1995.

[174] Thurber, C. Ritsema, J., Theory and Observations – Seismic Tomography and Inverse Methods, in Volume 1 of *Treatise of Geophysics*, edited by B.A. Romanowicz & A.M. Dziewonski, Elsevier, pp. 323–360, 2007.

[175] Trampert, J. & Woodhouse, J.H., High resolution global phase velocity distributions, *Geophys. Res. Lett.* , **23**, 21–24, 1996.

[176] Trampert, J., Vacher, P., & Vlaar, N. Sensitivities of seismic velocities to temperature, pressure and composition in the lower mantle. *Phys. Earth Planet. Inter.* , **124**, 255–267, 2001.

[177] Turcotte, D.L. & Oxburgh, E.R., Finite amplitude convective cells and continental drift, *J. Fluid Mech*, **28**(1), 29–42, 1967.

[178] Tushingham, A.M. & W.R. Peltier, ICE-3G: A new global model of late Pleistocene deglaciation based upon geophysical predictions of postglacial relative sea level change, *J. Geophys. Res.* , **96**, 4497–4523, 1991.

[179] Wahr, J., The forced mutations of an elliptical, rotating elastic and oceanless Earth, *Geophys. J. R. Astron. Soc.* , **64**, 705–727, 1981.

[180] Wang, Y. & Weidner, D.J., $(\partial\mu/\partial T)_P$ of the lower mantle. *Pure Appl. Geophys.* **146**, 533–549, 1996.

[181] Weertman, J., Creep laws for the mantle of the Earth, *Philos. Trans. R. Soc. London A* , **288**, 9–26, 1978.

[182] Weertman, J. & Weertman, J.R., High temperature creep of rock and mantle viscosity, *Annu. Rev. Earth Planet. Sci.* , **3**, 293–315, 1975.

[183] Woodhouse, J.H. & A.M. Dziewonski, Mapping the upper mantle: three–dimensional modeling of earth structure by inversion of seismic waveforms, *J. Geophys. Res.* , **89**, 5953–5986, 1984.

[184] Woodward, R.L. & Masters, G., Global upper mantle structure from long-period differential travel times, *J. Geophys. Res.* , **96**, 6351–6377, 1991.

[185] Wu, P. & Peltier, W. R., Glacial isostatic adjustment and the free air gravity anomaly as a constraint on deep mantle viscosity, *Geophys. J. R. Astron. Soc.* , **74**, 377–449, 1983.

[186] Xu, W., Lithgow-Bertelloni, C., Stixrude, L.& Ritsema, J., The effect of bulk composition and temperature on mantle seismic structure, *Earth Planet. Sci. Lett.* , **275**, 70–79, 2008.

[187] Zhang, S. & Christensen, U., Some effects of lateral viscosity variations on geoid and surface velocities induced by density anomalies in the mantle, *Geophys. J. Int.* , **114**, 531–547, 1993.

[188] Zhang, S. & Yuen, D.A. The influences of lower mantle viscosity stratification on 3D spherical-shell mantle convection, *Earth Planet. Sci. Lett.*, **132**, 157–166, 1995.

[189] Zhang, J. & Weidner, D.J., Thermal equation of state of aluminum-enriched silicate perovskite. *Science* , **284**, 782–784, 1999.